Event Related Potential Measures of Task Switching in the Implicit Association Test

Mark Andrew Coates, BSc.

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
In partial fulfillment of the requirements
For the PhD degree in Psychology

School of Psychology
Faculty of Social Sciences
University of Ottawa

Abstract

Since its creation in 1998, the Implicit Association Test (IAT) has become a commonly used measure in social psychology and related fields of research. Studies of the cognitive processes involved in the IAT are necessary to establish the validity of this measure and to suggest further refinements to its use and interpretation. The current thesis used ERPs to study cognitive processes associated with the IAT. The first experiment found significant differences in P300 amplitude in the Congruent and Incongruent conditions, which were interpreted as a reflection of greater equivocation in the Incongruent condition. The second experiment tested the task-set switching account of the IAT in much greater detail by analyzing each trial type separately. In the Congruent condition, all trial types elicited the same amplitude P300. Local probability, and the consequent checking and updating of working memory, was thought to be responsible for differences between trials of the Incongruent condition that required or did not require a task switch. The final experiment examined the role of working memory in the IAT by introducing obtrusive and irrelevant auditory stimuli. The results of Experiment 3 indicated that the introduction of an obtrusive and irrelevant auditory increment deviant has little overall effect on the IAT, and a similar effect on switch and no-switch trials within the Incongruent condition. This could have been because both the Congruent and Incongruent conditions of the IAT make such extensive demands on central processing resources that few are available to allow for the switching of attention, or it is possible that the IAT does not require significant updating of working memory. The usefulness of ERPs in the study of the IAT effect is demonstrated by the current research. In particular, the finding that behavioural results were not always consistent with the ERP results demonstrates that electrophysiological measures can complement traditional behavioural measures.

Table of Contents

List of Tables

List of Figures

List of Appendices

A Note on Formatting:

This dissertation is presented in article format. Therefore, the study that has already been published in a peer-reviewed scientific journal (Experiment 1: Coates & Campbell, 2010) is formatted and organized according to the specifications of the journal in which it was published, Neuroreport. As the primary investigator and main author of this study, M. Coates is listed as the first author on this publication. Other parts of the thesis, including the General Introduction, Pilot Study, Experiments 2 & 3, General Discussion, and reference section following the formatting guidelines of the American Psychological Association.

Chapter 1

Introduction

Traditionally, attitudes have been thought to include three components; a cognitive component (thoughts), an affective component (feelings), and a behavioural component (Fazio & Olson, 2003). Assuming each of these components to be within the conscious awareness of every individual, measuring attitudes could be done with a combination of self-report measures and behavioural observation. Of course, this assumption is suspect, as people are often not completely aware of their thoughts and feelings related to attitudes. Furthermore, even if assuming that one has access to their attitudes, individuals are often not willing to honestly share these with others.

Research into socially sensitive topics has revealed a dissociation between the behavioural and cognitive components of attitudes. For example, the classic study by LaPierre (1934) revealed that hotel owners tended to state a particular belief regarding how they would approach an interracial interaction (cognitive component) but behave differently (behavioural component) when they were actually confronted with that situation. It is likely that the social acceptability of certain responses led participants to respond to self-report measures in a way that is inconsistent with their actual attitudes. Similarly, studies have found that while most White individuals overtly report little or no racial prejudice on questionnaires, many show prejudiced behaviour (Schuman, Steeh & Bobo, 1997). It is likely that many individuals respond to explicit questions in a way that is socially acceptable but inconsistent with their actual attitudes.

Many (if not most) cognitive processes occur outside the level of conscious awareness. Many processes that occur at a non- or pre-conscious level can nevertheless greatly influence behaviour. With respect to attitudes, it is likely that non-conscious attitudes influence the overt expression of all three components of attitudes. However, without direct measurement of the implicit attitude, it is not possible to determine its relative contribution to the explicitly measured behaviour.

The Implicit Association Test (IAT) was designed as a measure of individual differences in implicit cognition (Greenwald, McGhee, and Schwartz, 1998). Specifically, the IAT sought to quantify and measure individual differences in implicit attitudes, which are thought to be outside of conscious awareness (the participant cannot access them) or attitudes which are within conscious awareness but socially unacceptable (the participant chooses not to report them). It was thought that the ability to measure attitudes without the confounding influence of self-presentation strategies would be useful to social psychologists that study topics where participants are unlikely to give completely accurate responses. Since its creation in 1998, the IAT has become a popular tool for measuring implicit attitudes due to its versatility, low-cost, large effect sizes, high face validity, and relatively high resistance to conscious manipulation. Issues surrounding the construct validity of the IAT still remain unresolved, especially in terms of the cognitive processes that underlie the task (Nosek, Greenwald, & Banaji, 2007). The purpose of this thesis is to use neuroscience method, Event Related Potentials, (ERPs) to monitor the cognitive processes associated with the IAT effect.

The IAT procedure

The IAT requires participants to quickly and accurately categorize stimuli (single words) presented serially on a computer screen. Stimuli fall into one of four categories; two attribute categories and two target categories. The subject typically pressed one of two keys to indicate the categorization of the stimulus. Each category consists of 5 to 25 individual stimuli (Nosek, Greenwald, & Banaji, 2005). The attribute categories are usually “good” and “bad”, with the “good” category including items such as “peace”, “friends”, and “love”, and the bad category including items such as “death”, “hatred”, and “disease”. The target categories used depend on the specific research question. For example, by using the target categories “male” and “female” the IAT could be used as a measure of sexism. The “male” category would consist of words such as “man”, “boy”, and “him”, while the female category would consist of words such as “woman”, “girl”, and “her”.

The word stimuli are presented to the participants as single trials in different conditions. In each condition, subjects receive a different set of instructions for how to respond. The two conditions of most interest are the "Congruent" and "Incongruent" conditions. In IAT research, "Congruent" refers to the condition in which words from pairs of categories (one attribute and one target category) that are generally associated with one another are paired on a single response key, and words from the other two categories are paired on the other response key. Using the sexism IAT described above as an example, in the “Congruent” condition the participants are asked to respond with one key when words in the “male” (“man”, “boy”) or “good” (“pleasant”, “happy”) categories are presented, and with the other key when words in the “female” (“woman, “girl”) or “bad” (“hate”, “agony”) categories are presented. Assuming that a participant was highly prejudiced, this task would be found to be relatively easy because the categories are mapped onto response keys in a congruent manner (for a sexist individual, “male” would be strongly associated with “good”, and “female” associated with “bad”). In a separate "Incongruent" condition, the categories “good” and “female” are now paired on the same key, while the “bad” and “male” categories are paired together on the other key. In the case of a sexist participant, the Incongruent condition would be much more difficult, since the pairings of good/female and bad/male are incongruent. In the sexism example of the IAT, RT differences between the Congruent (man/good; woman/bad) and Incongruent (man/bad; woman/good) would be explained by an implicit sexist attitude, that male words are more highly associated with positive words. In various studies of the IAT effect, RTs in the Incongruent condition are typically 100-200 ms longer than RTs in the Congruent condition. This difference in RT is defined as the IAT effect.

By changing the target categories, the IAT can be altered to be a measure of various types of implicit attitudes. For example, attitudes towards homosexuality can be assessed by using target words such as “gay”, “homosexual”, “heterosexual”, and “straight”, while attitudes towards different races could be assessed using stereotypical names from various racial groups.

Most IAT studies follow a procedure that is illustrated in Table 1. The task is divided into 5 conditions. The first condition is known as the *target acquisition condition*. Participants simply respond to words from one of the target categories (words such as “boy” or “man”) with the left key, and to words from the other target category (such as “girl” or “woman”) with the right key. In the second condition, called the *attribute acquisition condition*, participants respond to words from the “good” category (words such as “heaven” or “family”) with the left key, and to words from the “bad” category (such as “hatred” or “death”) with the right key. Thus, across the two conditions, one response key links the target category “maleness” (words such as “boy”, “man”) to the attribute “good” (“heaven”, “family”) while another response key links the target category “femaleness” (“girl”, “woman”) with the attribute “bad” (“hatred”, “death”). The third condition is the first combined condition, often called the *Congruent condition*. In this condition, words from one target category (male words such as “boy” or “man”) and from one attribute category (good words such as “heaven” or “family”) require a response with the same key (left), while words from the other target category (female words such as “girl” or “woman”) and the other attribute category (bad words such as “hatred” or “death”) both require a response from the other key (right). Typically participants find the Congruent condition to be fairly easy to complete quickly and accurately. In the fourth condition, the *target reversal condition*, only words from the two target categories are presented. The participant must now respond with the opposite keys from what they did in the target acquisition condition. For example, they would now respond to “male” words such as “boy” or “man” with the right key and would respond to “female” words such as “girl” or “woman” with the left key. In the fifth condition (known as the *Incongruent condition*) all four categories are again presented. Subjects must now respond to words from one target category (female words such as “girl” or “woman”) and from one attribute category (good words such as “heaven” or “family”) with the left key, while words from the other target category (male words such as “boy” or “man”) and the other attribute category (bad words such as “hatred” or “death”) require a response from the right key.

In most IAT paradigms, the order of presentation of the Congruent and Incongruent conditions has an effect on RTs in each of the combined conditions. That is, when the Congruent condition is presented before the Incongruent condition, RT differences between the two are generally greater than when the Incongruent condition is presented first. However, more recent research (Nosek et al, 2005) has indicated that increasing the number of trials in the practice conditions (target acquisition, target reversal) leads to more reliable results in the Congruent and Incongruent conditions, regardless of the order of presentation of the two combined conditions. The exact number of trials in each condition varies across studies. In the initial study by Greenwald and colleagues (1998), 20 stimuli were presented in each of the practice conditions. The Congruent and Incongruent conditions contained 80 trials each, of which the first 20 were used only for practice.

The IAT "effect" is defined as the difference in mean reaction time between the Incongruent condition and the Congruent condition. In the original IAT study (Greenwald et al, 1998) RTs in the Congruent condition ranged from around 500 to 700 ms, while RT in the Incongruent condition ranged from around 800 to 900 ms. In that study, the actual extent of RT differences between Congruent and Incongruent conditions varied depending on the target concept and the specific procedure used. For example, using an "insect/flowers" IAT, the average RT difference was between 113 ms when the Incongruent condition was presented first and around 213 ms when the Congruent condition was presented first. On the other hand, using an IAT measuring relative preference for Black and White faces, average RT differences ranged from 120 ms (Incongruent first) to around 198 ms (Congruent first), but order effects were not statistically significant. An improved scoring algorithm has been introduced which uses transformed RT rather than the raw RT (Greenwald, Nosek & Banaji, 2003). This scoring algorithm was developed in order to reduce artifact associated with RT and to maximize correlations between the IAT and self-report measures. Although the use of this scoring algorithm is essential for researchers interpreting the magnitude of the IAT effect in individual participants, the current research is not directly concerned with interpreting the scores of individuals, but rather with the cognitive processes involved in the complete task itself. Therefore the updated scoring algorithm will not be used in this research. More importantly, the present thesis also records the brain's responses (the so-called event related potentials or ERPs) that occur following the presentation of each stimulus. The mean of these single trials is then computed. The single trial ERPs are however not transformed. The interpretation of the mean RT based on transformed data and the mean ERP based on untransformed data would thus be difficult.

It is important to note that the IAT is considered to be a measure of relative strength of associations. For example, a Black/White IAT cannot not tell a researcher how much a particular participant dislikes Black people, only the relative preference that the participant has for Whites over Blacks. In this case, a significant IAT effect could be due to strong positive associations for White people but no associations for Black people, or strong negative associations towards Black people but no positive for White people. Other implicit measures have been developed to attempt to address this issue and allow for the assessment of attitudes towards a single construct (for example, Nosek & Banaji, 2001).

Since its creation in 1998, the IAT has been used to assess attitudes in several areas of psychology. Most common are studies involving attitudes towards various groups. Because the IAT can be customized to suit the specific questions of the researchers, it has been used to assess attitudes towards many different groups. Such topics include racism (Dasgupta, McGee, Greenwald, & Banaji, 2000; Hugenberg & Bodenhausen, 2003; McConnell & Leibold, 2001), sexism (Aidman & Carroll, 2003; Rudman & Glick, 2001), homophobia (Steffens & Buchner, 2003), and even relative preference for meat and vegetables in vegetarians (Barnes-Holmes, Murtagh, Barnes-Holmes & Stewart, 2006). The IAT has also been used to document changing attitudes at a population level. For example, Plant and colleagues (2009) reported that since the election of Barack Obama, there has been a significant change in the racial preference IAT effect among Americans, suggesting that this population holds a more positive (or less negative) view of African-Americans than in the past. In actual fact, at the time of the Plant and colleagues (2009) study, which took place in the fall of 2008, the IAT revealed no evidence of anti-Black bias at all. This is particularly meaningful when compared to previous research (for example, Nosek et al, 2007), which found over 60% of participants showing a bias towards White rather than Black faces.

Beyond its use in social psychology, the IAT has recently been used as a measure of consumer attitudes and preferences (Maison, Greenwald, & Bruin, 2004). In this study, the authors found that the use of the IAT predicted actual consumer behaviour (such as buying behaviour) beyond that which was predicted by explicit measures alone. This is somewhat surprising, since one would not expect consumer attitudes to be vulnerable to the same self-reporting biases that are present in measures of more socially sensitive topics such as racism.

What are implicit attitudes?

In the context of the IAT, implicit attitudes refer to attitudes that are outside of conscious control and awareness. There is some debate among researchers about the relationship between implicit attitudes, as measured by the IAT, and explicit attitudes or behaviour. Researchers have observed significant differences between explicit and implicit attitudes, even when the categories of interest were not socially sensitive, such as relative preference for flowers or insects (Greenwald et al., 1998). There are several possible interpretations of this finding. It is possible that a self-presentation bias results in participants responding to self-report questionnaires in a socially acceptable manner. A participant's lack of introspective access to their true attitudes could also confound the relationship between implicit and explicit measures. It is also possible that implicit attitudes could be an entirely separate theoretical construct from explicit attitudes.

A meta-analysis of 126 studies comparing the implicit measures (not limited to the IAT) to explicit measures (Hoffman, Gawronski, Gschwendner, Le, & Schmitt, 2005) found the correlation between explicit and implicit measures across studies to be moderate but significant ($r = 0.24$). The authors found that certain moderator variables accounted for a significant proportion of the variance. Implicit-explicit correlations were higher when the explicit self-report measure relied more on a participant's "gut-feeling" in order to make a response, and when the explicit measure assessed relative preference between targets (making it conceptually more similar to the IAT). The authors also found that correlations were higher when the explicit measure assessed affective responses (for example, "how do you feel about group X?"), rather than cognitive responses (for example, "do you associate certain traits with group X?").

Interestingly, the relationship between implicit and explicit measures has been shown to be malleable. A recent study (Koole, Govorun, Cheng & Galluchi, 2009) revealed that formal meditation promoted greater congruence of self-esteem as measured by implicit and explicit measures. This was thought to be related to meditation's promotion of greater use of intuitive feelings of self-esteem in the explicit measures. That is, following meditation training participants responded more quickly to explicit measures, and their responses were more consistent with implicit measures. While this study did not utilize the IAT, similarity between implicit measures suggests that the IAT may also be susceptible to outside influences. These types of concerns have led to a great deal of study into the fakeability of the IAT (Nosek, Greenwald & Banaji, 2007). Results have consistently shown that, compared to self-report measures, the IAT is less influenced by self-presentation bias. The IAT appears not be be influenced by vague instructions (for example: "try not to appear racist"), although explicit instructions can be given to participants to bring IAT performance under some degree of conscious control.

Relationship between implicit attitudes and behaviour

In addition to the debate over the relationship between implicit and explicit attitudes, there is much debate over the relationship between implicit attitudes and overt behaviour. Several authors have questioned the degree to which implicit attitudes affect behaviour. For example, McConnell and Leibold (2001) investigated the correlation between IAT scores, explicit measures, and observable behaviour during interracial interaction. Participants first completed explicit measures of racial bias, followed by a Black versus White IAT, and two brief interviews (one each with a Black and White experimenter). For the explicit measures, participants were asked to rate how they felt towards Blacks on a scale of 0 to 100 (with 0 being least favourable and 100 being most favourable), and then completed the same task for their attitudes towards Whites. Subjects were also given a 7-point scale with different attributes at each end of the scale (such as “beautiful/ugly”, “honest/dishonest”, and “good/bad”), and asked to score each race for each of these attribute pairs. The correlation between explicit measures and the IAT was found to be moderate ($r=0.42$). The interviews were also scored to determine differences in the amount of certain meaningful social behaviours such as eye contact, and smiling. Moderate correlations were found between IAT scores and several of the social behaviours, including speaking time ($r = 0.51$), smiling ($r = 0.39$), speech errors ($r = 0.42$) and speech hesitation ($r = 0.35$). In contrast, attitudes measured by explicit questionnaires were not significantly correlated with any of the social behaviours measured. The authors interpreted this as evidence that implicit attitudes, as measured by the IAT, were related to non-conscious behavioural biases in interracial social interactions (McConnell & Leibold, 2001).

A study by Richeson and Shelton (2003) found that the IAT effect predicted the degree to which an individual’s performance on a Stroop task (naming the colour of ink that words are written in, while inhibiting the tendency to read the words) would be impaired during interracial interactions. That is, White individuals who scored in a more biased fashion on a Black versus White IAT showed a greater interference effect on a Stroop task when tested by a Black examiner, while the IAT effect was not related to Stroop performance when tested by a White examiner. This study demonstrates that implicit attitudes can predict observable behavioural differences.

Egloff and Schmukle (2002) were interested in the ability of the IAT to predict behaviour related to anxiety. They modified the IAT slightly to be a relative measure of the associations between the concepts of self ("me", "I") versus others ("him" "her") and anxious ("worry", "tremble") versus calm ("peace", "serenity"). In a series of experiments, the authors found that IAT anxiety scores were significantly correlated with non-verbal behaviour while performing a stressful task. Furthermore, using a hierarchical regression model with the IAT score being entered as the third and final step, IAT score accounted for a significant proportion of the variance beyond that accounted for by self-report measures. The authors interpreted this as strong evidence that the IAT could be used as an implicit predictor of anxiety and the resulting behaviour.

Asendorpf, Banse, and Mucke (2002) took this reasoning a step further by differentiating between spontaneous behaviour and controlled, non-spontaneous behaviour. The authors had their participants engage in a videotaped conversation with an attractive, unfamiliar, confederate of the opposite sex from the participant. Behavioural indices of shyness were divided into spontaneous behaviour, which included facial and body movements and tense body posture, and controlled behaviour, which included speech and demonstrative movements. The participants then completed an implicit shyness IAT, as well as explicit self-report measures of shyness. Results indicated that IAT scores were significantly correlated with spontaneous shy behaviour, while self-report measures were significantly correlated with controlled shy behaviour. This double dissociation provides some of the strongest evidence yet that implicit attitudes are a separate construct from explicit attitudes, and that both can have an influence on behaviour.

In a review article of IAT, Nosek, Greenwald and Banaji (2007) reported that discriminatory behaviour is more strongly correlated with the IAT than with self-report measures. Both explicit measures and the IAT were significantly correlated with consumer choices and political preferences, but self-report measures had a slightly higher correlation than the IAT. This is consistent with the assumption that the IAT measures attitudes that are either inaccessible or deliberately withheld; while discriminatory behaviour is unlikely to be reported on a self-report measure, there is somewhat less reason to believe that consumer preference or political affiliation would be consciously withheld.

Controversy over cognitive processes involved

The IAT has been largely developed in the context of the study of implicit attitudes in the field of experimental social psychology. Thus, many of the studies examine the extent to which the IAT effect reflects or correlates with observable attitudinal bias. Very few studies have actually examined the cognitive processes that could account for the difference in RT between Congruent and Incongruent conditions. Cognitive psychologists have long employed the speed of decision making, as indexed by RT, to measure hypothetical steps, or "stages" of information processing. Longer RTs are presumed to reflect either additional stages of processing or additional time with a stage or stages of processing.

A significant gap in the literature still exists with respect to the cognitive processes that underlie the IAT. The original explanation of the IAT (Greenwald, McGhee, and Schwartz, 1998) rested on the assumption that the Congruent task would simply be "easier" to perform than the Incongruent task, if the participant were highly prejudiced against one of the categories. Thus, RT would be faster if the words from the category “good” (such as “peace”, or “health”) were paired on the same response key as the words from the category “male” (words such as “man”, and “his”). Similarly, the RT would be faster if the words from the category “bad” (such as “death”, or “sadness”) were paired on the same response key as the words from the category “female” (words such as “woman”, and “her”). While this explanation is certainly face valid, the original article did not make any attempt to explain exactly *why* the Congruent task would be easier, or what cognitive processes made it “easier”. Although the original creators of the IAT did not investigate the cognitive processes involved, they have since indicated that this is an important area for future research:

“Having a more comprehensive model of [cognitive processes involved in the IAT] would likely benefit research using the IAT, possibly by affording greater efficiencies in administration, by suggesting design changes that would increase the IAT’s construct validity as a measure of association strength, or by making further progress toward developing effective non-relative measures of association strength. ” (Nosek, Greenwald, Banaji, 2007)

Furthermore, in a recent review of implicit measures such as the IAT, (De Houwer, Teige-Mocigemba, Spruyt, & Moors, 2009) it was argued that understanding the cognitive processing giving rise to the observable behaviour in these tests is of the utmost importance. These authors defined an implicit measure as one that results in observable behaviour that is caused automatically by psychological attributes. That is, the results of the measure are not related to any behavioural goals of the individual in question. Therefore, in order to establish that a given task is an implicit measure, it must be demonstrated that performance on the task is determined (as least in part) by automatic processes. The authors argue that greater knowledge of the cognitive processes involved in a task can help to establish the construct validity of the IAT and other implicit measures.

A measure may lack construct validity if factors other than those being studied have a significant effect on task performance. For example, if individual differences in speed of information processing, which presumably is not related to attitudes, has a reliable effect on RT in the IAT, the IAT could be said to be measuring something other than attitudes. The issue of construct validity is crucially important because the IAT has been used as a dependent variable to assessment implicit attitudes in hundreds of studies. If the IAT is a valid measure of the construct of implicit attitudes or bias, then these studies will have provided useful insight into these constructs. However, if the IAT lacks construct validity and is actually measuring different cognitive processes, the conclusions of many of these studies could be called into question. Understanding the cognitive processes involved during the IAT would have significant implications for its validity as a measure of implicit associations.

A limited number of cognitive explanations of the IAT effect have been offered by past researchers (see for example, Brendl, Markman, & Messner, 2001; De Houwer, 2001; Mierke, & Klauer, 2001; Rothermund & Wentura, 2001, 2004; Conrey, Sherman, Gawronski, Hugenberg & Groom, 2005). Each of these theories will be described briefly below.

Rothermund and Wentura (2001) propose a model for understanding the IAT that differs from the original explanation provided by Greenwald, McGhee, and Schwartz (1998). The authors argue that rather than measuring relative strength of *associations* between concepts, the IAT actually measures congruence of what they call *salience* of the stimuli. By this account, when two salient categories are mapped to the same response key it is relatively easier to perform the task than when a salient category is paired to the same key as a less salient category. In this context, salience is defined as a property of the stimulus or category that causes it to “stand out” from the other stimuli being presented. For example, a picture of the face of someone from a dominant cultural group would not be particularly salient, while the face of a visible minority would stand out more. Presumably, stimuli that are unusual, novel, disturbing, or unknown would be more salient than well-known, recognizable figures. Consider a race IAT using the categories “good”, “bad”, “White”, and “Black”. For an individual living in a society in which White people are the majority, the stimuli associated with the category “Black” would stand out due to the relative infrequency of exposure to Black people. Similarly, stimuli from the “bad” category should also stand out due the fact that they are relatively more disturbing, and possibly more biologically meaningful than pleasant words. Thus, on a typical IAT task when the key pairings are “good/White” and “bad/Black” the categories on a given key are said to be Congruent by Greenwald and colleagues (1998). Rothermund and Wentura (2001) argue that these pairings have similar salience. When the key pairings are “good/Black and bad/White” the categories mapped to a given key are said to be Incongruent by Greenwald and colleagues (1998), and that this accounts for the increased RT. Rothermund and Wentura (2001), however, claim that the pairings have different salience, and that this leads to the increased RT. Thus, two categories may appear to be highly associated with one another, while in fact the only common property that they share is their salience.

Although Rothermund and Wentura’s interpretation of the IAT effect is interesting, it does not directly answer the question of what cognitive processes can explain the IAT effect, it merely replaces “associations” with “salience” in the model originally proposed by Greenwald (1998). In fact, Greenwald, Nosek, Banaji, and Klauer (2005) have argued that this account is little more than a different semantic interpretation and offers minimal insight into the mechanisms involved in the IAT. Furthermore, it is possible that the tasks used in these studies involve very different cognitive processes from those involved in the standard IAT despite similar RT findings. RT differences can emerge for many different reasons. They need not be the same in the usual IAT procedure and that employed by Rothermund and Wentura. Furthermore, some authors (Greenwald et al., 2005) have criticized the salience asymmetry account based on its limited applicability across IAT paradigms. In a replication of the original Rothermund and Wentura article, these authors found that salience was not related to performance on a gender identity IAT, despite obtaining robust IAT effects.

De Houwer (2001) argued that the cognitive processes involved in the IAT are similar to those involved in other stimulus-response compatibility tests such as the Stroop task. In these types of test, performance is enhanced (lower RT) when the response required is similar to some characteristics of the stimulus. For example, suppose a participant were instructed to respond with left and right key presses to arrows pointing either left or right. Participants will respond faster when the the required response matches the spatial characteristics of the stimuli (press the left key when presented with a left facing arrow) than when the instructions require a stimulus—response incompatibility (press the left key when a right facing arrow is seen).

The IAT is conceptually similar to this task due to the short-term, learned associations between positive and negative valence and the response keys. That is, a participant learns in the Congruent condition to associate "good" words with the right hand key. When the Incongruent condition is completed, they are instructed to inhibit this learned association, creating a stimulus-response incompatibility. Participants might process the valence of the individual target stimulus, or they might process the valence of the category to which a stimulus belongs. These learned associations may explain why the IAT effect is sometimes greater when the Congruent condition is presented before the Incongruent condition.

In previous studies of the IAT, all individual target stimuli had a valence that was either neutral or consistent with the valence of the category itself. For example, in a racism IAT, each individual exemplar of the Black and White categories would be a neutral face or name, without any specific valence associations. In De Houwer's experiment (2001), the valence of individual target stimuli was varied such that it was either consistent or inconsistent with the valence of the category as a whole. For example, while the target category “British” was considered to have positive valence for the participants, individual exemplars were varied to include pictures of Britons who tended to elicit either positive (Princess Diana) or negative (Margaret Thatcher) reactions in most individuals. Similarly, while the category “Foreign” was generally associated with a negative valence, exemplars of foreign individuals could have either a positive (Mahatma Ghandi) or negative (Saddam Hussein) valence. The results of this experiment indicated that only the valence of the target category had an effect on performance. That is, the positive or negative valence of individual exemplars was not related to any difference in RT. Thus, RT to, for example, “Gandhi” was long because of the association with “foreign” while RT to “Thatcher” was short because of the association with “British”. This study provided important information about the stimulus characteristics that are relevant to the IAT effect. Although this study appears to provide little information about the cognitive processes involved in the IAT, it does demonstrate that the characteristics of the exemplars related to the target categories are more important in determining IAT effect than the valence of the individual stimuli.

Brendl, Markman, and Messner (2001) have suggested an interpretation of the IAT effect that does describe the cognitive processes involved in the task. These authors argue that shifting “response criteria” contributes to the IAT effect. According to these authors, participants use different response criteria in the Congruent and Incongruent conditions. On any given trial, information from the stimulus is extracted until a decision can be reached about which response is necessary. Attribute stimuli (for example words such as “happy” in the “good” category) contained only valence information, while target stimuli (for example “boy”) contain both identity information (“male” category) and valence information (positive associations, in the case of a sexist participant). While the identity information is needed to correctly respond to target stimuli, the valence information is not needed and in fact must be ignored for correct performance in the Incongruent condition. The authors used a random walk model to describe the decision making processes. In the random walk model, the participant starts each trial at a neutral point then is pushed towards one decision or the other by the gradual accumulation of information extracted from the stimulus. According to this model there are two ways that RT could be longer in the Incongruent condition than the Congruent condition. First, the valence and identity information could compete with one another, thus slowing down responses to the target stimuli (but leaving responses to attribute stimuli unchanged). Alternatively, as the authors argue, more information needs to be extracted (a “higher threshold”) before a response is initiated in the Incongruent compared to the Congruent condition. This requirement of additional information will also delay RT. The authors called this this shifting response criteria hypothesis.

Brendl, Markman, and Messner (2001) tested the two hypotheses in an experiment employing a modified IAT which used non-words as one of the target categories and White names as the other target category. Since attitudes towards non-words would presumably be neutral, the authors claimed that prejudice or attitudes could not contribute to the IAT effect seen in this paradigm. The result of the study were in line with the predictions made by the shifting response criteria hypothesis. That is, RT of all four categories (White names, non-words, good words and bad words) were increased in the Incongruent condition. Although this study provides an interesting way of conceptualizing the decision making process used in the IAT, it suffers from one of the same limitations as the studies of Rothermund and Wentura (2001, 2004). Non-words may not be processed in the same way as real words, and the claim that non-words are perceived as neutral was left untested. Furthermore, the authors make the assumption that the decision making process on each trial begins at a neutral point, with no influence from the preceding trials. This assumption is unlikely to be supported, as sequential effects have been observed in many cognitive paradigms. Mierke and Klauer (2001) described the “task switching” theory, which provides a more satisfying explanation by taking into account the effect that previous cognitive task sets have on performance on subsequent trials.

Conrey, Sherman, Gawronski, Hugenberg and Groom (2005) proposed a multinomial processing tree model of implicit association test performance. According to this so-called “Quad Model”, decisions are made at multiple stages of processing, each decision being dependent on the preceding decisions. According to the model proposed by Conrey and colleagues (2005), performance on each trial of the IAT is influenced by the probability that an automatic bias is activated by the stimulus, that the correct response to the trial can be determined, and that automatic, biased responses can be withheld. Finally, if an appropriate response is not determined, the model predicts that subjects will guess based on information not conveyed by the stimulus. While the Quad model holds promise for improved understanding of the IAT, it can only be evaluated by examining the error rates in various conditions. Differences in RT across conditions is the basis of the IAT effect, and any model that does not address RT differences cannot fully explain the task as it is used by researchers in the real world.

Mierke and Klauer (2001, 2003) argue that the difference in RT between the

Congruent and Incongruent conditions is the result of “task switching” costs. The IAT can be conceptualized as two separate tasks, an attribute categorization task and a target categorization task. On any given trial, the participant is required to either categorize an attribute stimuli (for example, Good versus Bad, the attribute categorization task) or categorize a target stimuli (for example, White versus Black, the target categorization task). Thus, the participant must first evaluate the stimulus to decide which of the two tasks must be performed, and then must determine which response is needed.

When two categories that are highly associated with one another (for example, “good” and “musical instrument”, or “bad” and “weapon”) are paired on the same key (the Congruent condition), the participants does not actually have to switch between these two tasks because a heuristic shortcut allows the task to be performed correctly. The valence information of the target stimuli (eg: "Musical Instrument" = “good”) can be used to determine the correct response. As such, a participant can correctly complete the task without categorizing the targets at all, simply by relying on the valence information contained in any individual stimulus. Rather than determining whether the attribute categorization task or the target categorization task needs to be performed, a participant can simply respond based on the valence for all trials; the valence of each stimulus provides sufficient information to determine the correct response. In the Incongruent condition, on the other hand, a participant is not able to rely on a single decision. The participant must categorize the stimulus as either an attribute (Good or Bad) or a target (Music or Weapon). As such, it takes the participant additional time to switch back and forth between the attribute categorization task and the target categorization task. The authors suggest that the IAT effect is therefore a function of the extra time required to switch between two different tasks in the Incongruent condition.

A series of experiments (Mierke & Klauer, 2001) was run to test this hypothesis. The authors observed that on Incongruent condition trials immediately following a task set switch (eg: an attribute category item followed by a target category item), RT was significantly longer than on trials that did not follow a task set switch (mean of 1033 ms compared to 860 ms). RT was also slower on trials that required a response switch, and RT was fastest overall when neither a task switch nor a response switch were required. In the Congruent condition, which supposedly can be completed without switching between the two tasks, task switching costs were non-significant.

In a second study, Mierke and Klauer (2001) experimentally manipulated task switching costs in order to examine whether task switching was causally related to the magnitude of the IAT effect. The authors provided cues prior to the target items. Cues were pictures depicting either one of the two target categories (flowers or insects) or one of the two attribute categories (thumbs up or thumbs down). The cue was presented on the right or the left hand side of the screen, according to the correct response for that category. These cues did not provide specific information about the category, valence, or correct response for the stimulus that would follow on the individual trial. It simply reminded the participant of the correct response to one of the four categories. That is, being shown the “flower” cue did not necessarily indicate that the subsequent item would be a flower item, or that it would be a positive valence item, or whether the correct response would be the left or right button (although, since the cues were randomized, there was always a certain chance that the cue would signify the correct response or up-coming category). As predicted, providing cues to participants reduced the IAT effect compared to an IAT without cues. The authors interpreted this finding as evidence that providing cues reduced task switching costs and thus reduced the IAT effect.

A subsequent study by Klauer and Mierke (2005) provided further evidence in support of the task switching account of the IAT. While their first set of studies investigated the most obvious predictions of the task set switching model, such as whether or not differences existed on any given trial based on what had preceded it, this more recent study examined at a less obvious prediction made by the theory. Specifically, the researchers were interested in carryover effects, also referred to as “task-set inertia”. Once engaged in a particular task, it is more difficult for a participant to switch to a new task. Conversely, if a given task must be suppressed, it is subsequently more difficult to activate. The authors believed that this may help explain the order effect in IAT studies (a larger IAT effect is observed when the Congruent condition is performed before the Incongruent condition rather than when the Incongruent condition is performed first). According to this hypothesis, when the Incongruent condition is completed before the Congruent condition, it becomes more difficult to activate the attributes (“good” or “bad”) of a target stimulus in the Congruent condition, resulting in slower performance. When the Congruent condition is presented first, on the other hand, participants readily adopt the strategy of relying on attributes (good or bad) for all stimuli.

In this experiment, Congruent or Incongruent conditions of an IAT task (insect/flower) and a new condition (a so called “test block”) were alternated twelve times. In the test blocks, the participants were required to indicate the attribute (good/bad) of target stimuli (insect/flower). To respond correctly in the Incongruent condition, the participants would need to ignore or suppress the attribute information (good/bad). If this suppression of attribute information carried over into the test block, task-set inertia would result in slower performance in the subsequent test block. In a Congruent condition, on the other hand, participants do not have to suppress the attribute information of the target stimuli. One group did all Congruent IATs, the other did all Incongruent. The results were in line with their expectations; the group that performed the Incongruent condition of the IAT performed more poorly (longer RT and more incorrect categorizations) on the test block than the participants who completed the Congruent condition of the IAT. This experiment demonstrated that participants can become “stuck” on one method of responding and presumably would need to use additional processing resources in order to switch back and forth between tasks.

A recent study by Klauer, Schmitz, Teige-Mocigemba, and Voss (2009) employed a latent variable approach to determine the specific aspects of cognitive functioning that are related to performance on the IAT effect. The authors used a battery of neuropsychological tests designed to capture several different aspects of executive functioning. Measures were used to assess task set switching (for example, alternately adding and subtracting from a given number), inhibition (for example, the Stop Signal task, a measure similar to the Go-NoGo task), and working memory capacity (for example, evaluating the veracity of sentences while retaining in memory the last word of each sentence). Multiple versions of the IAT were administered in order to extract a latent variable representing underlying IAT performance (which was presumably due to method specific variance) rather than attitudes towards the specific categories used in each task. Results indicated that overall IAT performance was more highly correlated with task set switching measures ($r = -.65$) than with measures of working memory ($r = -.31$) or inhibition ($r = -.04$). A model without the path between task set switching and IAT performance lost a significant amount of goodness-of-fit, supporting the notion that task set switching is independently related to method variance in the IAT to a greater extent than either working memory capacity or inhibitory ability. Although working memory ability was examined, the authors minimized its role in the IAT effect. When the contribution of inhibition was removed from the model, working memory capacity was related to IAT performance over and above the effects of task set switching. That is, capacity for inhibition and working memory may both have played an important role in IAT performance, but accounted for overlapping variance. It is surprising that the authors did not discuss the role of working memory in detail, but instead choose to focus alone on task switching ability as the cognitive trait related to IAT performance.

This study provides strong support for the notion that task set switching and working memory are both critically important in the IAT. However, it is limited by the fact that it is a correlational study. It is possible that there exists a third factor that is related to both IAT performance and task set switching. In the title of the paper, the authors refer to “flexibility”, presumably as a cognitive or personality trait which allows greater ability to switch between task sets. Unfortunately, this is never again mentioned in the text of the article.

The task-set switching account is currently considered to be one of the strongest theories describing the IAT effect (Nosek, Greenwald & Banaji, 2007). The implications of the theory are consistent with the general interpretation of the IAT. However, interpretation of the IAT rests on its validity as a measure of implicit attitudes. The possibility that the IAT could be lacking in construct validity is highlighted by its use of RT as the primary measurement. Cognitive psychology has long noted that many different factors can affect RT. Cognitive operations have a past, a present and a future. Past cognitive activity might have had a subsequential effect on the present response, as is apparent in the Mierke and Klauer (2001) interpretation of trial-to-trial variation and task switching on the IAT effect. The present response might have consequential effects on future activity, for example, RTs could be slower after the participant commits an error or after switching from one task set to another. Unfortunately, the participant’s current response does not provide an overt measure of the cognitive events that led to the response nor its consequential effects. Rather, these cognitive processes must be *inferred* on the basis of RT.

Thus, there is much debate about the nature and interaction of the hypothetical "stages" of processing leading to the overt response. RT differences may reflect the difference in the time required to classify stimuli, to select an appropriate response to a particular stimulus, or to implement that response (McCarthy & Donchin, 1981). This thesis specifically examines the factors that are thought to account for this significant RT difference between Incongruent and Congruent conditions. The present set of studies therefore will employ ERP methodology to complement the more traditional use of the behavioural RT in an attempt to isolate the differences in cognitive processes responsible for the IAT effect.

Social neuroscience

Social neuroscience is an emerging field which seeks to utilize the insights and methodologies of neuroscience to deepen our understanding of social psychological and social cognitive phenomena (Mitchell, 2008). Technologies such as functional magnetic resonance imaging (fMRI), which allows oxygen utilization in the brain to be measured during various activities, have been used to study brain areas involved in social cognition. Social neuroscience research has been criticized in the past on the basis that it provides few new insights into social cognitive processes, beyond simply providing colourful images of brain regions activated during various cognitive tasks. While it is true that some of the published studies in this field have reported empirical findings without contributing significantly to our theoretical understanding of the phenomenon in question, other studies have shown this approach to be a valuable way of refining theories which were based on evidence from behavioural measures alone. Some of these experiments have involved phenomena related to the IAT, such as prejudice or bias.

A recent meta-analysis (Van Overwalle, 2008) examined the brain regions involved in two main types of social cognition. Combining the data from over 100 fMRI studies, the author discriminated between two main forms of social cognition with different neurological substrates: judging the mental state of other individuals, and making judgments about members of a group based on past experience with the group. Judging the mental state of others (so called “mind-reading”) was found to activate a network involving the temporal parietal junction. This system is thought to be based on perception of temporary behaviours which allows an individual to infer to mental state of another person, sometimes referred to as a “mirror system”. In contrast, the medial prefrontal cortex was found to be most involved when participants were required to base decisions on semantic knowledge about an entire group or category of people. That is, when making inferences about people of a particular group or category, activity in the medial prefrontal cortex is related to the accessing of knowledge based on previously learned associations about that group of people. Because results from social neuroscience experiments were consistent with what was previously known of these brain regions, researchers are able to infer some of the cognitive processes involved in social phenomena.

An fMRI study by Mitchell, Ames, Jenkins and Banaji (2009) provided further evidence that relying on stereotypes for decision-making involves the areas of the brain that are also used more generally for categorization of information. The authors conducted fMRI scans while participants read a series of statements that were either gender stereotypical (“enjoys shopping for clothes” or “enjoys action movies”) or neutral (“likes Coke more than Pepsi” or “enjoys drinking coffee in the morning”). During the presentation of each sentence, a picture of either a man or a woman was also displayed on a monitor. Notably, all of the statements related to preferences of the hypothetical person, rather than physical attributes or occupations. fMRI revealed greater activation in the middle and inferior right frontal lobes when stereotypes were activated, compared to neutral trials. Furthermore, the extent of activation in right frontal areas was significantly correlated with performance on a gender IAT. That is, those participants who strongly associate gender with traditional gender roles showed greater activation in the right frontal lobe when making decisions in a stereotyped manner, compared to participants with more egalitarian beliefs.

The authors of this study note that the right frontal lobe has previously been associated with numerous functions, including inhibition of inappropriate behaviour, categorization and semantic retrieval. The authors state it was unlikely that inhibitory processes resulted in the observed differences in frontal activation, since those participants who had greatest activity in this area were the ones who appeared less able to inhibit stereotypical responses. Of course, it could be argued that these were the very individuals who were unable to inhibit behaviour (or expressions of "attitudes") that society would consider to be inappropriate. It is also possible that the fMRI differences reflect an effect of semantic congruity/incongruity associated with traditional gender roles. Thus, the authors argued that participants make judgments about other people's mental state on the basis of pre-existing beliefs about an entire category of people. This process is assumed to be subserved by the right frontal lobe, which is also involved in accessing other semantic knowledge. Stereotypical beliefs about or actions towards others, therefore, may be due to the over-reliance on this general approach to estimating the beliefs of others.

In addition to the research on explicit social cognitive processing described above, a fair amount of research has been devoted to the role of distinct areas of the brain in implicit social cognition. For example, Cunningham and colleagues (2004) found that distinct brain regions are involved in processing faces of different races. In White subjects greater activation in the amygdala, an area of the brain thought to relate to fear and emotional processing, was elicited following presentation of Black faces compared to White faces. This effect was most apparent when stimuli were presented for a relatively brief period of time (30 ms). The difference in amygdala activation was decreased when faces were presented for over 500 ms, presumably due to voluntary control of emotional processes. Furthermore, Black faces elicited additional activation in the dorsolateral prefrontal cortex, the anterior cingulate, and the ventrolateral prefrontal cortex, areas thought to be involved in executive control of cognitive activity.

A disadvantage of fMRI is that while it provides a wonderful window to view the brain regions involved in various activities, it provides little insight into actual sequencing (or ordering) of processing. A major reason for this is that the fMRI reflects changes in oxygenation of the different brain regions during the experimental manipulation compared to a neutral baseline condition. Oxygen is of course carried by the circulating blood and thus ultimately fMRI reflection variation in blood flow to the different brain regions. Blood flows much more slowly than the rate at which many cognitive processes take place. fMRI is therefore unable to detect very rapid changes in brain activity. This limits its usefulness for investigating cognitive processes that occur very rapidly.

Brain regions associated with the IAT

While a fairly large amount of social neuroscience research has examined implicit attitudes, this thesis is one of the first attempts to describe the electrophysiological underpinnings of the IAT. As there is relatively little information available from which to draw theoretically based predictions, interpretation of the ERP findings is likely to represent the greatest challenge in this research. Previous studies using cognitive neuroscience methods are limited to one lesion study (Milne & Grafman, 2001), and two studies conducted using fMRI (Chee et al., 2000; Phelps et al., 2000). Very recently, researchers using ERP have begun to investigate the IAT effect (Hurtado et al, 2009; Williams & Themansen, 2010).

Studies of patients with brain lesions provide useful information about the brain areas involved in specific cognitive processes. Presumably, damage to an area involved in a certain cognitive process will cause a deficit in the ability to perform that process. The result could be a complete inability to perform tasks requiring that process, or a decline in some measure of performance (such as error rate or RT). An important criticism of lesion studies is that tasks typically require the use of multiple cognitive processes. For example, a task such as the Stroop test (Stroop, 1936) requires processes involved in reading, colour perception, inhibition, and speech production. Damage to areas involved in any of these processes could result in impaired performance on this task. Nevertheless, lesion studies provide an excellent starting point for isolating the brain areas involved in the cognitive tasks, and maybe provide insights into the cognitive processes involved.

The ventral prefrontal cortex is thought to be involved in the storage of social knowledge that the brain uses to make appropriate decisions in social situations (Milne & Grafman, 2001; Grossman et al, 2010). Damage to this area is therefore linked with socially inappropriate behaviour, as seen in patients with ventral lesions (Stuss, Gow, & Hetherington, 1992; Adolphs, 2009). A study (Milne & Grafman, 2001) of ten patients with ventromedial prefrontal lesions found a reduced IAT effect (95 ms) when compared to normal controls (208 ms), or patients with frontal lobe lesions in areas other than the ventral prefrontal cortex (340 ms). The authors argued that stereotypes, a form of implicit social knowledge that presumably are stored in the ventral prefrontal cortex, are no longer accessible by these patients, resulting in the elimination of the IAT effect. It could also be speculated that damage to the dorsolateral prefrontal cortex, an area known to be involved in inhibitory processes, was related to the greater IAT effect in patients with frontal lesion not in the ventral prefrontal cortex. One major limitation of lesion studies is the heterogeneity of the sample used in the study. That is, none of the subjects in this study had damage in the exact same areas of the brain, and many had suffered damage to surrounding brain areas. However, this study provides evidence which suggests that the ventral prefrontal cortex is involved in retaining or accessing the information that is used by participants during the IAT.

In a more recent study, Gozzi and colleagues (2009) examined performance on a Male/Female IAT in a much larger sample of brain injured patients. Due to the much larger sample size (154 patients), this study had greater power to detect differences between more specific sites of injury. After controlling for the size of the lesion, the researchers found that patients with damage to either the ventromedial prefrontal cortex or the anterior temporal lobe (another area thought to be involved in the storage of social knowledge) showed an increased IAT effect relative to normal controls. In contrast, patients with damage to the ventrolateral prefrontal cortex had a diminished IAT effect relative to normal controls. The ventromedial prefrontal cortex has been linked to the control of emotional reactions, and lesions in this area may therefore alter a patient's ability to respond in a socially desirable fashion on the IAT. Likewise, the anterior temporal lobe (which includes the amygdala) is known to be involved in the processing of fear and other emotions. Damage to this area could therefore reduce emotional responses associated with social judgment or stereotypes. The authors argue that ventrolateral prefrontal cortex lesions could have affected IAT performance through their role in executive processes such as inhibition, although why damage to this area resulted in improved performance on the Incongruent condition (which presumably is more difficult) compared to the Congruent condition is unclear. Participants were also administered a second IAT task, with the target categories of "Vietnam" and "hockey" (the participants were Vietnam veterans). Performance on this task was not associated with damage to any of the studied brain areas, suggesting that the areas in question are specifically related to processing of social attitudes and stereotypes, rather than the structure of the IAT itself.

Other neuroscientists have employed fMRI to image brain areas that are thought to be involved with attitudes. The amygdala, a part of the limbic system with connections to the frontal lobes and the hippocampus, is thought to be involved in emotional learning, particularly fear conditioning. Phelps and colleagues (2000) used fMRI to measure amygdala activation when participants looked at White and Black faces. Participants then completed several direct and indirect measures of racial prejudice, including the IAT. The authors argued that some participants have a negative emotional evaluation of Blacks, which would be expressed by greater activation in the amygdala when looking at pictures of faces of Blacks. The authors found that greater amygdala activation was correlated with a greater IAT effect. The study did not however employ fMRI during the actual IAT. Nevertheless, this study does suggest that both the IAT effect and differential amygdala activation are related to the learned negative associations that develop through individual experience.

Chee and colleagues (2000) conducted a study using fMRI during the actual administration of the IAT. These authors reasoned that the IAT effect might be explained by differences in inhibitory processes needed to complete the two conditions. That is, during the Congruent task, participants do not have to inhibit responses consistent with a previously learned association between categories (good and music, or bad and weapons, for example). However, there is a tendency to respond according to these previously-learned associations in the Incongruent task as well. That tendency must be inhibited if the task is to be performed correctly. The need to inhibit the automatic response and the need to activate the correct response results in slowing of RT. In this study, increased activation of the left dorsolateral prefrontal cortex, a region previously linked to inhibitory processes, was observed in the Incongruent task relative to the Congruent task. However, it has also been shown in other research to be involved in working memory (Petrides, 2000) and memory encoding (Murray & Ranganath, 2007). When there is no way to distinguish between the various potential processes that could be involved in the task, information about the timing and duration of cognitive operations is arguably more useful than simply knowing the regions of the brain that are involved.

Taken together, the above results suggest involvement of at least three major brain areas, the dorsolateral prefrontal cortex, the ventromedial prefrontal cortex, and the amygdala. Studies using fMRI and patients with lesions certainly contribute to our understanding of the brain areas involved in the IAT. However, from the perspective of functional operations, they provide little information about the specific timing of the cognitive processes involved in this task. fMRI images represent changes in blood flow. Because blood flows slowly (in the order of seconds), the methodology has relatively poor temporal resolution. RT in the IAT is in the order of 400-800 ms, much faster than fMRI is capable of measuring. A methodology with better temporal resolution than fMRI would have numerous advantages over fMRI.

Event related potentials

Event-related potentials are the minute changes in the ongoing electrical activity of the brain (the EEG) that are elicited by an external stimulus or an internal psychological event. A major difficulty with the use of ERPs is that they are represent such small variations in the ongoing EEG that they are difficult to distinguish from background noise. Signal averaging across numerous trials is a technique which allows the meaningful ERPs to be separated from the background variations in the EEG. Using this technique, one makes the assumption that the signal (ERPs) is constant with repeated presentations of the stimulus while the background noise is random. If there are a sufficient number of trials, the average amplitude of the background noise should approach zero, which would allow the constant ERPs to emerge. The background noise cannot be completely eliminated, but its amplitude is reduced as a square root function of the number of trials (that is, to reduce the noise by a factor of two would require four times the number of trials). The background noise can also be reduced by filtering specific frequencies in the EEG, but this risks removing any actual ERP that overlaps this frequency (Campbell, 2010).

ERPs consist of a series of positive and negative components. The different components can be categorized according to the time (or "latency") at which the component reaches maximal amplitude. For example, the P3 category of components, which is associated with active discrimination tasks, typically peaks 300 ms or longer after stimulus onset. The P3 component can be further subdivided into sub-components, including the earlier passively elicited P3a, and P3b (often referred to as P300).

Components of the ERP can also be categorized according to the experimental manipulations that affect them. Components that are mainly affected by physical aspects of the stimuli (intensity, frequency, etc) and are minimally affected by psychological processes such as attention are referred to as exogenous ERPs. On the other hand, components that are mainly affected by the psychological relevance of the stimulus are called cognitive or endogenous ERPs. The longer latency components such as P300 typically reflect endogenous processes.

The ERP components recorded from the scalp represent the volume-conducted activity of different sources within the brain. This allows ERP components to be distinguished according to their intracranial sources. The ERP waveform is made up of numerous components, each of which is generated in a specific area of the brain and elicited by a specific process. The overall peaks of the ERP waveform can be directly measured through scalp recording, but the individual components, and the intracranial sources of those components, must be inferred from the results of experimental manipulation of the stimuli or cognitive processes involved. Two ERP peaks that have different scalp distributions must necessarily have different intracranial sources.

Background noise, described above, is not the only potential artifact in the EEG. A phenomenon known as “latency jitter” can change the appearance of the actual peaks due to variance in peak latency. This is particularly true in experiments with small numbers of participants or individual trials (Möcks, 1986). Variance in latency of a peak is indirectly proportional to peak amplitude (Johnson, 1986); the more variable the latency across subjects and individual trials, the smaller than amplitude of the peak will appear in the grand average. This is especially problematic for experiments in which there is a large difference in variance of latency between two or more conditions; this can result in the peak in one condition appearing smaller due to being being “smeared” and the peak in the other condition appearing larger, even if both conditions elicit peaks that appear the same size in the grand average.

Two components of the ERP of of primary interest to the current thesis, the N2 and the P300. The N2 component is a negative going deflection peaking at approximately 180 ms when elicited by visual stimuli. The N2 is thought to comprise several subcomponents, each reflecting different cognitive processes. An early frontal-central N2 is involved in mismatch or novelty detection, while a later frontal-central N2 is thought to reflect cognitive control processes such as response inhibition, strategic monitoring, response conflict monitoring, and error monitoring. Posterior N2 components are related to shifts of visual attention (Folstein & Van Petten, 2008). For the current research, the most relevant subcomponent is the second frontal-central N2 involved in cognitive control. The Incongruent condition of the IAT is thought to be difficult in part because the participants must inhibit the tendency to respond to highly associated target and attribute words with the same key. Novelty detection is unlikely to affect N2 amplitude in the IAT, since individual stimuli are presented an equal number of times throughout each condition.

Paradigms that are known to elicit this subcomponent include tasks that require the suppression of a prepotent response. The most commonly used such task is the so-called "Go-NoGo" task, in which a participant must respond to most stimuli, then withhold a response when a certain rare stimulus is presented. In other words, the paradigm causes the participant to become biased towards a certain response, then requires that response to be withheld. The N2 elicited in such a paradigm is larger on "NoGo" trials (withholding a response) than on "Go" trials (on which the participant makes the typical response). The N2 is thought to reflect not just inhibitory processes involved in withholding the prepotent response, but also in response conflict monitoring and other aspects of cognitive control (Hammerer, Li, Muller, & Lindenberger, 2010). Cognitive control would be necessary on any trials that require the participant to respond in a way that is different from their prepotent tendency.

The P300 component of the ERP was discovered in the 1960s. Since that time, researchers have developed a strong understanding of the characteristics of the P300, and the conditions that affect it. Specifically, the latency and amplitude of this component have been used as an index of the timing and nature of cognitive operations. This knowledge allows researchers to make inferences about which cognitive processes are involved in a particular task. For example, the cognitive processes involved in the well-known Stroop task (Stroop, 1936) which is though to require inhibition of the prepotent tendency to read words in favour of naming the dissonant ink colour, have been evaluated using ERP.

Several studies been conducted on the temporal characteristics of brain activity in the Stroop task (Stroop, 1936) in an attempt to understand the cognitive processes involved. This task involves two conditions. In both conditions, words (names of colours) are presented in coloured ink. In the word-relevant condition participants are instructed to simply read a word presented to them. This is sometimes referred to as the Congruent task, since reading words is Congruent with the automatic response made by most people when presented with a word. In the colour-relevant condition, the participants are instructed to ignore the word and say the name of the colour in which the word is presented. For example, if the word "Green" were presented in the colour yellow, the participant would have the say "Yellow". This is sometimes referred to as the Incongruent task, since saying the name of the colour requires the participant to inhibit the automatic tendency to read the name of the colour. Duncan-Johnson and Kopell (1981) were interested in the characteristics of the P300 component of the ERP, which is thought to reflect stimulus evaluation processes (Kutas, McCarthy, & Donchin, 1977; Duncan-Johnson & Donchin, 1982). They argued that if the delay on the colour word interference task was due to greater stimulus evaluation time, the latency of P300 should increase in conjunction with RT. On the other hand, if delayed response selection accounts for the increased RT, the P300 should remain the same across all trials. Results indicated that P300 latency was nearly constant across the word-relevant (Congruent) and colour-relevant (Incongruent) tasks. This supports the hypothesis that the delay in RT on the Incongruent task is the result of increased response selection processes, such as response competition.

The Incongruent block of the IAT is conceptually similar to the Incongruent block of the Stroop task. In both tasks, stimuli are presented sequentially and the subject must withhold their prepotent response. That is, in the usual Stroop task the participant must inhibit the automatic tendency to read the words in order to correctly name its colour. In the IAT, the participant must inhibit the automatic tendency to respond to semantically similar categories with the same key. Both paradigms result in a slowing of RT compared to a Congruent condition in which participants respond in a more automatic fashion. The use of ERPs with the IAT may allow for conclusions to be drawn about the cognitive events that contribute to RT differences between Congruent and Incongruent blocks, in the same way that processes underlying the Stroop test were investigated.

Although P300 latency was thought in the past to be related only to stimulus evaluation processes (Kutas, McCarthy & Donchin, 1977), more recent research has indicated that when participants are performing a task quickly, both stimulus evaluation and processing of responses can affect latency (Verleger, 1997). It has been argued that P300 reflects some cognitive process involved in both perception of a stimulus and response initiation (Verleger, Jaskowski, & Wascher, 2005). To test this hypothesis, latency of response-locked P300 peaks were compared with latency of stimulus-locked peaks. P300 amplitude was found to be equally large in both cases, suggesting that both stimulus evaluation and response processes were involved. Therefore, P300 latency may not provide a sensitive measure of whether the IAT effect is related to stimulus evaluation, or to decision making processes. However, P300 latency may still provide some useful information about the IAT effect. Kok (2001) reported that peak latency is longer when more processing resources are required.

Peak latency is also affected by individual differences (Polich, 2007). For example, superior mental performance, particularly on measures that involve rapid processing of information, is correlated with shorter P300 latency. Changes over the lifespan have also been noted. Latency decreases as children age, then increases throughout adulthood. Further increases in peak latency are observed in individuals with dementia or other neurological conditions affecting the brain, as well as some psychiatric disorder such as depression. Situational factors can also affect latency, such as caffeine intake or consumption of other drugs. Such potential confounds will likely not be problematic in IAT studies, since all experiments will be within-subjects designs and therefore will remove the variance associated with individual differences.

The triarchic model of P300 amplitude

Johnson (1986) argues that variations in the amplitude of the P300 component of the ERP can be explained by the combined effects of three dimensions: the subjective probability of the stimulus, the meaning of the stimulus, and the information transmitted by the stimulus. Each of these dimensions can be further broken down into more specific components. Subjective probability can be broken down into a priori probability and sequential probability. Stimulus meaning can be subdivided into stimulus complexity, stimulus value, and task complexity. Finally, information transmission can be divided into the effects of equivocation and attention. Subjective probability and meaning are thought to have additive effects on amplitude, while information transmission has a multiplicative relationship with the other two dimensions.

Subjective stimulus probability is probably the best understood of the three dimensions. The dimension of subjective probability is thought to operate automatically, without the need for conscious regulation. Subjective probability can be subdivided into two components, a priori probability and sequential probability. There is a clear inverse relationship between the **a priori probability** of a stimulus and the amplitude of P300; more rare stimuli evoke a larger amplitude P300. This effect has been shown in multiple types of paradigms, regardless of whether the participant is instructed to count stimuli, react quickly to oddballs, or receives feedback on performance (Johnson, 1986). Since all stimuli in the IAT are presented the same number of times, it is unlikely that a priori probability has an effect of P300 amplitude in this paradigm.

In addition to the a priori probability of the stimulus, P300 amplitude is affected by the **sequential expectancy** of the participant. That is, participants come to expect that a certain sequence of stimuli will be presented, and violations of that sequence cause a larger P300 amplitude. The first example of this phenomenon (Squires et al., 1976) showed that repeated stimuli elicited a smaller amplitude P300 than a non-repeated stimulus. Participants expected each stimulus to be repeated, and this violation of sequential expectations was related to increased P300 amplitude. In the IAT, stimuli are presented in a random order, which should minimize sequential expectancy effects. However, even a random sequence will contain strings of similar stimuli. It is possible that P300 is elicited differently during such a string of stimuli than it would be following presentation of a subsequent item that changes the pattern. For example, suppose five stimuli in a row were all from the category “weapons”, then the sixth stimulus was a “good” word. The sixth stimulus would be expected to elicit a greater amplitude P300 than the fifth stimulus, since it would represent a violation of the expectancy created by the preceding trials. The potential role of sequential expectancy will be discussed in the section on the task-set switching account of the IAT effect.

Stimulus meaning has appeal as an explanation of the P300 differences observed in the IAT. After all, the IAT itself was developed to detect associations between concepts or categories, which implies that the meaning of each stimulus (at least in relation to other categories) is being evaluated on each trial of the task. In other words, some characteristic of the stimulus needs to be evaluated to complete the task correctly. Stimulus meaning has three components, as described by Johnson (1986): task complexity, stimulus complexity, and stimulus value.

Task complexity may be the aspect of stimulus meaning that provides the strongest basis for interpreting the P300 amplitude effects in the IAT. Task complexity is argued to have a direct relationship with P300 amplitude, such that more complex tasks result in larger amplitude P300 following stimulus presentation. By all accounts the Incongruent task is much more complex than the Congruent task, and this complexity results in the longer RT in this condition. The model predicts that P300 will be larger in the more complex Incongruent condition than in the simpler Congruent condition. While this prediction is fairly clear it is not particularly helpful. The original IAT article by Greenwald and colleagues (1998) described the Incongruent condition as being more difficult than the Congruent condition, but did not provide any deeper insight into the cognitive processes involved. The same would appear to be true about the effect of task complexity on P300.

The model by Johnson (1986) predicts that more complex stimuli elicit larger amplitude P300, independently of task complexity. **Stimulus complexity** is unlikely to play a role in the IAT effect. It would be difficult to argue that any of the stimuli are more or less complex than others. All stimuli were fairly common words, were roughly the same length (1-3 syllables), and were presented in the same font. It is possible that certain words were more or less accessible in memory, which would cause those words to require different levels of processing prior to response selection. However, this effect would be unlikely to last beyond the first few presentations of any particular stimulus. Regardless, it would not explain any differences between Congruent and Incongruent conditions, since the exact same words in the same sequence are presented in the two conditions.

Stimulus value is a component of the stimulus meaning dimension that has appeal as an explanation for the amplitude differences observed, in part because it relates to one of the previously described theories to account for the IAT effect (Rothermund & Wentura, 2001, 2004). As discussed above, valence processing is thought to occur automatically outside of conscious awareness. Bad words could arguably be considered to have higher value than good words, since it would be adaptive for words signifying potential danger to draw attention to a greater degree than words that do not signal danger (neutral or good words). This would be consistent with the salience account of the IAT effect proposed by Rothermund and Wentura (2001). Bad words and weapons would be more salient and draw attention to themselves more than good words and musical instruments, and this model would therefore predict larger amplitude P300 on the more salient trials of all conditions. While stimulus value would perhaps explain differences in P300 amplitude elicited by the various categories of stimuli (if they were to be analyzed separately), it would not provide any explanation for the differences between the Congruent and Incongruent conditions, since the exact same stimuli are used in both conditions.

Information transmission is the third main dimension thought to affect P300 amplitude. Information that can be conveyed by a stimulus includes information about subsequent trials or about responses to be made. Information transmission is closely related to task performance. In fact, Johnson (1986) states that any change in P300 amplitude accompanied by a change in task performance necessarily indicates a difference in information transmission. Since there are clear performance differences (RT and error rate) between the Congruent and Incongruent conditions of the IAT, there is a strong rationale for suspecting that a different amount of information is being conveyed in the two conditions. There are two components of the information transmission dimension: the amount of information conveyed by the stimulus that is lost due to the participant's uncertainty about whether they correctly perceived the trial, and amount of information lost due to lack of attention.

Loss of information due to perceptual uncertainty is known as **equivocation**. Experiments have shown increased certainty that a stimulus was perceived correctly to be directly related in P300 amplitude and behavioural performance. Notably, even when several stimuli are designed to contain the same amount of information, the amount of information that participants chose to use can affect performance and P300 amplitude. For example, in the original study to demonstrate the effects of equivocation (Johnson & Donchin, 1978) participants were required to estimate the passage of one second, then received feedback from one of two tones which indicated whether they were correct or incorrect. The volume difference between the two tones was varied, which led to a difference in perceptual certainty (when the tones were more similar to one another, it was more difficult for a subject to tell which tone had been presented). In this task, P300 amplitude and task performance were significantly lower when the tones were more similar to one another. When participants were simply told to count the tones, however, there was no significant difference in P300 amplitude or performance. The same amount of information was conveyed by the tones in both tasks, but in the time estimation task the subjects could also choose to rely on their experience with the task in order to perform it correctly. The option to use internal information depended on the difficulty of the time estimation task, such that subjects relied less on the feedback cue when it was more difficult to perceive. Therefore, the authors concluded that when a stimulus is difficult to perceive, subjects have the option to reduce their use of information from the stimulus itself, relying instead on internal cues. This strategy is related to both a performance decrease and P300 amplitude reduction.

The other dimension of information transmission that affects P300 amplitude is the amount of **information lost due to inattention**. For example, if a participant is told to ignore a series of tones, less information will be transmitted by those stimuli than by the same series of tones when they are attended. Research has shown that attended stimuli produce larger amplitude P300 than unattended stimuli. The stimuli are the exact same in the Congruent and Incongruent condition, but one could argue that different characteristics of some of the stimuli must be attended in the two conditions. For the attribute categories (good and bad words) the same characteristic of the word (valence) must be assessed in order to provide the correct response in both the Congruent and Incongruent condition. However, for the target words (weapons and musical instruments), the valence of the word is sufficient to respond correctly only in the Congruent condition. In the Incongruent condition the category of the word must be processed and the valence information must be ignored. This could lead to an overall lower P300 amplitude in the Incongruent compared to the Congruent condition.

It should also be cautioned, however, that P300 latency is related to both RT and P300 amplitude. Johnson (1986) points out that P300 amplitude is inversely related to variability in latency (more variability in P300 latency is related to smaller P300 amplitude in grand averages). It is possible that peak latency in the Congruent and Incongruent conditions will differ in their level of variability. For example, suppose the Congruent task produced less variable peak latency. This could result in an artificially greater apparent amplitude of P300.

Therefore peak latency will need to be examined in detail before conclusions can be drawn about the effects on P300 amplitude.

Double peaks and double decisions

Since P300 is thought to reflect decision making processes regarding the stimulus or response, a double peaked P300 may reflect two decisions being made. Many studies have used ERP to detect multiple decision making using a variety of different paradigms. In a study conducted by Johnson and Donchin (1985), subjects were provided with feedback about their accuracy on a task requiring them to estimate time. A double peaked positivity (P300) was elicited by the feedback stimulus. The first peak was interpreted as a reflection the classification of the feedback as right or wrong. The second peak was though to reflect an updating of working memory and possible correction of error as a result of the feedback. This study demonstrated that P300 could provide a measure of internal (endogenous) cognitive processes. Thus, multiple cognitive events can be measured using ERP, even in cases where behavioural measures such as RT cannot distinguish between the various cognitive events taking place during a given task.

Muller-Gass, Gauthier, Desrochers and Campbell (2001) presented native French speakers with a French noun on a computer monitor. Beneath the noun, the words “un” and “une” were presented on one condition and “masculine” and “feminine” in another. The subject’s task was to press the key appropriate to the gender of the noun. Thus, if the word “table” were presented, the subject would press the key associated with “une” in one condition and “feminine” in the other condition. When categorizing the gender of the stimuli according to "un" or "une", a single P3 peak was observed. However, there was a doubled peaked positivity when categorizing the gender according to superordinate labels. This suggests that a double decision is used by subjects when categorizing stimuli according to superordinate labels. The authors argue that the first decision reflects a decision about the gender of the word, while the second is thought to reflect the decision about which label ("masculine" or "feminine") to apply to that word.

ERP and social neuroscience/prejudice research

ERPs have also been successfully used to study a number of social cognitive processes, including those related to attitudes. For example, Bartholow, Riordan, Saults and Lust, (2009) examined the roles of strategic control and response conflict in an affective priming paradigm. Primes and targets were positive and negative English words. Participants were presented with a 200 ms duration prime, followed by a 50 ms blank screen and then a target stimulus. In different blocks of trials, the probability of affective congruence between the primes and targets was manipulated (20%, 50%, or 80% probability of congruence). Thus, all blocks were made up of individual trials which could be congruent or incongruent, but the probability of congruency varied across blocks. As expected, performance (error rate and reaction time) improved with greater congruency between primes and targets. The lateralized readiness potential (LRP) revealed response activation in the motor cortex after the onset of the priming stimulus. Notably, in the condition in which the prime was most likely to be affectively incongruent with the target, the prime elicited activation of the response incongruent with the prime. This indicated that participants were strategically expecting the target to be affectively different from the prime. In contrast, when the prime was more reliably congruent with the target, the prime caused participants to prepare to activate the response congruent with the prime. Similarly, a larger amplitude N2 component (thought to reflect response conflict monitoring) was elicited on incongruent trials when the probability of congruence was high, but N2 was smaller on incongruent trials when probability of congruence was low. P3 amplitude followed the same pattern, with a larger amplitude on incongruent trials in the high congruency block, and smaller amplitude on incongruent trials in the low congruency block. The authors argued that this experiment revealed that behavioural affective congruency effects are related to the actions of the response activation system. Furthermore, they claimed that P3 amplitude was enhanced when it was difficult to predict the target based on the prime, which suggests that the P3 reflects recognition regarding whether or not the current response strategy is correct. This study is relevant to the IAT because it is possible that strategic response monitoring and control differ between the Congruent and Incongruent conditions of the IAT in a similar fashion.

Ito and Bartholow (2009) reviewed electrophysiological correlates of race perception and processing. Several ERP components, both exogenous and endogenous, have been shown to be sensitive to racial status of stimuli. N100 and P200 are both larger in response to stimuli related to racial out-group members, possibly reflecting greater attention to relatively novel targets. N200, in contrast, is larger when viewing faces of in-group members. This may relate to great depth of processing of facial features of members of the participant's own race. P300 is larger when violations of stereotype expectations between primes and targets occurred (Bartholow, Dickter, & Sestir, 2006). While not directly related to the IAT itself, race-IATs may involve many of these same processes and elicit similar electrophysiological responses.

In another study conducted using ERP (He, Johnson, Dovidio & McCarthy, 2009) researchers found significant differences in the ERPs elicited by White, Black, and Asian faces, and revealed that these differences in processing were correlated with the magnitude of an individual's IAT effect. Several early visual components differed between conditions, which the authors suggest was related to visual processing of different physical characteristics of the stimuli (such as colour). Later cognitive ERPs also differed across stimuli. Specifically, the amplitude of P200 at frontal scalp sites differed depending on the race of the face (Black > White), while N200 amplitude showed the opposite pattern. A slow positive component (LPC) at T6 was also significantly larger following presentation of White faces compared to Black or Asian faces. Several ERP components were significantly correlated with participant's performance on racial IAT measures. Earlier components showed the greatest correlations between the Black/White IAT and ERP differences, while the largest correlation between Asian/White IAT and ERP differences was found at the later components. The authors argued that, compared to processing White faces, automatic or implicit processes are involved when seeing Black faces and more deliberative processes are involved when viewing Asian faces. The authors speculate that similar underlying brain regions (for example, activity in the amygdala, the fusiform gyrus, or frontal regions) may be involved during the IAT as during passive viewing of faces. However, in this study the IAT and the face viewing task were not completed at the same time, so there is no direct evidence that similar cognitive processes take place during the IAT as during the face viewing task.

ERPs and the IAT

The previously described studies of the electrophysiological correlates of prejudice demonstrate the usefulness of ERPs in the study of implicit cognition. However, they do not directly address the cognitive processes involved in the IAT. The IAT is often used as a measure of prejudice, but it does not follow that the processes involved in prejudice will necessarily be involved in the IAT. The IAT can be adapted as a measure of implicit preference for any pair of categories, including “self” versus “other” or “Coke” versus

“Pepsi”. It would be surprising if the exact same cognitive processes and neural circuitry involved in racial stereotyping was also involved in the implicit preference for soft drinks. Researchers have raised concerns regarding the variance in IAT performance that is due to process variables in addition to stereotypes or attitudes. Such process variables, or “method specific variance” would apply to all IATs, not just those concerned with prejudice. Therefore, a more general approach to understanding the IAT is needed, which provides information about the neural correlates of the task itself and not just the correlates of a construct that the IAT is used to measure.

Barnes-Holmes and colleagues (2004) provide the first published study of ERP and the IAT. This study was examining the IAT in the context of Relational Frame Theory (RFT), an operant conditioning model of language acquisition. The IAT type task used in this study was modified significantly from the standard version. Associations between pseudowords were learned within the experimental paradigm itself. That is, participants were taught that certain pairs of pseudowords are associated with one another. In the Incongruent condition, the authors observed a large negative peak occurring at approximately 1200 ms, which was maximally larger over left frontal scalp sites. In the Congruent condition, this peak had a latency of approximately 1300 ms and was positive. Unfortunately, because the methodology was so different from previous IAT studies and the results so different (extremely long RTs), interpretation of the ERP results in the context of the standard IAT paradigm is not possible.

The first of only two studies to examine ERPs associated with an actual IAT was conducted by Hurtado and colleagues (2009). This study compared ERPs in indigenous and non-indigenous Chilean participants performing a racial IAT. This study was mainly concerned with the frontal “Late Positive Potential” (LPP) elicited approximately 400 ms after stimulus onset. Differences in LPP amplitude (larger amplitude and more lateralized to the right hemisphere in the Incongruent condition) were found to be significantly correlated with IAT effect and self-reported prejudice. However, this result was only observed when stimuli were words and not when stimuli were faces, a somewhat puzzling finding given previous research which found differences in electrophysiological correlates of processing faces of different races. The authors argued that this finding was due to valence evaluation of words being more explicit than evaluation of faces. Similarly, the effect of condition on LPP was only noted in indigenous participants, a very surprising finding considering that participants from both groups showed behavioural evidence of prejudice.

This study was primarily concerned with the processing of words and faces with different valence in the context of the IAT. Other cognitive processes such as task switching or shifting response criteria that have been proposed as explanations for the IAT effect were not addressed. Another major concern about this study is that it employed a data-mining technique to determine electrophysiological measures that were correlated with behavioural and self-report measures. In effect, they first identified variables that were highly correlated with behavioural measures and then ran statistical analyses only on those variables, a technique that has been vigorously criticized in the context of fMRI research (Vul et al, 2008) for its tendency to artificially inflate correlations between imaging data and behavioural measures.

The study by Hurtado and colleagues (2009) makes an important contribution to the understanding of how valence of stimuli in the IAT are processed. Interestingly, however, it provides little insight into the differences between the Congruent and Incongruent conditions. Since stimuli are the exact same in both conditions but processing is clearly different (as evidenced by longer RT in the Incongruent condition), understanding of the role of the valence in the IAT provides only a first step in the understanding of the IAT effect.

A second study to examine the IAT using ERPs was recently published (Williams & Themansen, 2010). This exploratory study examined differences in various ERP components elicited by the Congruent and Incongruent conditions of a Homosexual-Heterosexual IAT. Components of interest included N1, P2, N2, N400, and LPP. It was predicted that early exogenous components (N1, P2) would not differ between conditions, while later components, which are thought to be associated with cognitive processes, would differ. Specifically, smaller N2 and N400, and larger LPP in the Incongruent condition were predicted.

No differences in N2 were observed between conditions, while the results for N400 and LPP were consistent with predictions. Differences in amplitude of N400 and LPP were significantly correlated with IAT effect (the difference in RT between Incongruent and Congruent conditions) in a curvilinear fashion, suggesting that processes reflected by these components may play a role in behavioural differences between the two conditions. Differences in N400 were thought to reflect semantic incongruency in the Incongruent condition, while LPP differences were interpreted as reflecting emotional incongruency. While this argument makes sense at the level of the entire condition (the categories paired to each key are incongruent with one another), on any given trial the participants are only responding to a single category. There is no incompatibility within each stimulus, only between the responses required for categories of stimuli. In addition, although the authors do not discuss P300 in this article, the scalp distribution and latency of the LPP resemble P300. Interpreting the differences in the LPP as differences in emotional congruency is only one of many possible interpretations of this component, particularly if it does reflect similar processes to those which elicit P300. The authors offered no theoretical explanation for their finding that the relationship between IAT effect and N400/LPP differences was curvilinear.

The authors suggest that a more thorough study of the effects of sequence on the IAT would be helpful in the understanding of the cognitive processes underlying the IAT effect.

There are two major methodological differences between the current procedures used in the current thesis and these two previous studies. First, the articles by Hurtado and colleagues (2009) and Williams and Themansen (2010) both employed versions of the IAT measuring bias or prejudice towards particular social groups. In the case of the Hurtado and colleagues study, an indigenous/non-indigenous IAT was used, while the Williams and Themansen study used a Gay/Straight IAT. It is not possible to determine with certainty whether the ERP findings in those studies were related to processes involved in the IAT itself, or to those related specifically to processing information about the social groups used as stimuli in the studies. This is of particular concern because previous research has shown that viewing pictures of members of social out-groups elicits different ERPs than viewing pictures of members of an in-group (He et al, 2009). To determine the processes involved in the IAT itself (rather than those related to processing information about members of a particular social group), the best version of IAT used should be one that includes target categories that do not represent specific social groups. Thus, the version of the IAT used in the experiments for the current thesis (weapons/musical instruments) was thought be a representative IAT, in that it would be minimally confounded by processes related to the specific target categories. The results of the current study may provide better evidence of the cognitive processes involved in the IAT, while the other two studies provide evidence of the processes involved in the specific versions of the IATs used in those studies.

The use of pictures as stimuli for the target categories and words as stimuli for attribute categories may be responsible for some of the differences between the results of the current thesis and the previous articles. Both studies found substantial differences between the ERPs elicited by the words and the pictures. In the study by Hurtado and colleagues (2009), this potential confound may be even more problematic due to the fact that words were presented for three times longer than were the pictures of faces (300 ms compared to 100 ms). It is unfortunately not possible to determine if this is due to differential processing of the categories themselves (ie: gay/straight versus good/bad), or whether this was due to differences in visual processing of the stimuli themselves. Certainly the early visual ERP components (N200, P200) are likely to differ across trial types, as previous research (He et al, 2009) has shown that different images can affect these components. Whether or not the later cognitive ERPs are affected by the modality of the stimuli (pictures or words) is not entirely clear. Although the current research does not suffer from this potential confound, it could be argued that the other two studies provide greater external validity, as many versions of the IAT employ pictures for the target words. Further research should compare the results of the two versions of the IAT (words for all categories compared to words for attributes and pictures for targets) using the same participants, in order to better understand the effects of stimulus modality.

The study by Hurtado and colleagues (2009) used a scoring algorithm suggested by Greenwald and colleagues (2003) which computes standard scores for each individual based on various manipulations of the reaction time and error rate of each participant. In contrast, the current research and the study by Williams and Themansen (2010) both used raw RT and error rate for all analyses. This decision was made in part because direct comparison of transformed RTs and raw ERPs is problematic. Furthermore, it is mainly studies in which individual differences between participants are being investigated that benefit from using the updated scoring algorithm. As the current research (as well as the study by Williams and Themansen) was interested mostly in the cognitive processes involved in the IAT and only employed RT as a secondary variable of interest, the use of this scoring algorithm was not considered to be necessary. In addition, the study by Hurtado and colleagues (2009) revealed that they obtained similar results with both scoring methods, suggesting that the use of the scoring algorithm would not have substantially change the findings of the current research.

Conclusion

Despite its widespread use and popularity as a measure of implicit attitudes, questions remain about the cognitive processes involved in the IAT. Previous studies have employed behavioural measures only, functional neuroimaging, and patient populations. Each of these approaches has significant limitations that do not allow the studies to fully address the relevant questions. Recently, studies have qualitatively examined ERP differences between conditions of the IAT. Both noted that future experiments should study specific cognitive theories of the IAT effect. The purpose of this dissertation is to use electrophysiological data from ERPs to test a cognitive theory that have been previous proposed as an explanation of the IAT effect.

The first study reported is a pilot study initially conducted to ensure that ERP methodology could be successfully used to identify differences between conditions of the IAT, and to provide the basis for future studies of this task. The second study is a follow-up and replication of the pilot study. The third study directly tests the task-switching hypothesis of the IAT by examining different types of individual trials. The final study of this thesis examines the role of working memory in the IAT by examining the effect of irrelevant auditory distractors on ERPs and behaviour.

Chapter 2

Pilot Study

Prior to directly evaluating specific theories of the cognitive processes involved in the IAT, it was necessary to demonstrate that ERPs could be successfully used in the study of the IAT. An exploratory pilot study was therefore conducted with the goal of qualitatively describing the ERPs elicited during the IAT. As this was the first such study using ERPs to investigate the IAT, there were no definite predictions prior to this study, except that the Congruent and Incongruent conditions were expected to elicit different ERPs. The specific nature of those differences was unknown at the time of the study, but would provide a starting point for future, theoretically driven experiments.

Participants

Participants were seven right handed university students. One participant was excluded from the analysis because frequent movement and blinking rendered the data unusable. Two of the six participants included in the analysis were female. None of the participants reported consuming unusual amounts of caffeine on the day of the experiment, and none reported a history of psychiatric illness.

Procedure

A modified IAT paradigm was administered. The target categories selected for this study were “weapons” and “musical instruments”. These specific target categories were chosen because they were among the categories used by previous researchers (for example:

Greenwald, McGhee, & Schwartz, 1998). Using these specific stimuli allowed for direct comparison of the current results with the behavioural results obtained in the past.

Stimuli consisted of four categories of words (Good, Bad, Musical Instruments, and Weapons), each of which consisted of 25 exemplars (see Appendix 1). The individual stimuli were selected from a previous IAT study (Greenwald et al, 1998). Items were presented in black letters on a white background at the center of a computer monitor at a viewing distance of approximately 50cm. Stimulus duration was 200 ms and the inter-stimulus interval (stimulus onset-to-onset) was 1200 ms. The first condition, known as the target acquisition condition, consisted of 25 trials from each of the two target categories (Musical Instruments and Weapons), for a total of 50 trials. Since there were 25 exemplars in each category, there was only one presentation of each exemplar. Similarly, the attribute acquisition condition consisted of 25 trials from the “good” category and 25 trials from the “bad” category, for a total of 50 trials. In the Congruent condition, 50 trials from each category were presented. Similarly, in the Incongruent condition there were also 50 trials from each category. Each individual exemplar was presented twice in each of the Congruent and Incongruent conditions. The Congruent and Incongruent conditions were each presented twice, although only the second condition was used in the analysis. This is consistent with previous IAT studies, in which practice trials are given to participants before those trials that are ultimately used in the analysis are presented. The second Congruent and Incongruent conditions were presented immediately following the first, and were in all ways identical. The order of the Congruent and Incongruent conditions was counterbalanced, such that half of the participants received the Congruent first and half received the Incongruent first. After the Congruent or Incongruent condition (whichever was received first), the subjects received the target reversal condition, which consisted of 25 trials from each of the two target categories (Musical Instruments and Weapons), for a total of fifty trials.

During the experiment, participants were asked to fixate on a central fixation point on a computer monitor, thus avoiding eye movement. In all conditions, participants were instructed to respond as quickly as possible while maintaining accuracy. Consistent with previously published studies using the IAT, participants responded to stimuli using one key on the left hand side of a keyboard and one key on the right hand side of a keyboard. In all conditions, participants used the index finger of the left hand to press the “a” key, and the index finger of the right hand to press the “l” key. Participants were instructed to keep their fingers on the keys between trials.

In the target acquisition condition, participants were instructed to press the “a” key when the word presented was a musical instrument, and the “l” key when the word presented was a weapon. In the attribute acquisition condition, participants were instructed to press the “a” key when the word presented was a good word, and the “l” key when the word presented was a bad word. In the Congruent condition, participants were instructed to press the “a” key when the word presented was either a musical instrument or a good word, and the “l” key when the word presented was either a weapon or a bad word. In the Incongruent condition, participants were instructed to press the “a” key when the word presented was either a weapon or a good word, and the “l” key when the word presented was either a musical instrument or a bad word. In the target reversal condition, participants were instructed to press the “l” key when the word presented was a musical instrument and to press the “a” key when the word was a weapon.

Electrical activity was continuously recorded from 5 scalp sites (Fz, Pz, Cz, F3, F4) and the 2 mastoids (M1, and M2) using silver-silver chloride electrodes attached to an elastic cap. The EEG was referenced to the tip of the nose. Vertical and horizontal electro-oculograms were also recorded. Inter-electrode impedances were below 5 kΩ. The high filter was set at 30 Hz and the time constant at 2 s Data were acquired continuously at a 256 Hz sampling rate. Offline, the data were subsequently filtered at 15 Hz using an inverse FFT algorithm.

Eye movement and blinking artifact were subsequently corrected also using an inverse FFT algorithm (Woestenberg, Verbatan, & Slangen, 1983). The data were later reconstructed into discrete trials beginning 100 ms before stimulus onset and continuing for 900 ms following it. Only trials in which a correct response was made were averaged.

Results

The average RT for the Congruent condition was 585 ms (*SD* = 145 ms), while the average RT for the Incongruent condition was 643 ms (*SD* = 187 ms). With this very small sample size, this difference was non-significant ($p < 0.6$). RT histograms (50 ms bins) are presented below the ERP results in Figure 1. In the Congruent condition the error rate was 0.15, while in the Incongruent condition the error rate was 0.33.

ERPs revealed several differences between the Congruent and Incongruent conditions of the IAT. In both conditions, a positive peak was apparent over frontal and central scalp sites (Fz and Cz) at about 280 ms following stimulus presentation. This peak was larger on the Incongruent condition than in the Congruent condition. At around 400 ms a negative peak, maximal over frontal sites, was observed. There was no apparent difference in the amplitude or latency of this peak between the Congruent and Incongruent condition. This is followed by a late positive component at roughly 500 ms following stimulus presentation during all trials of the IAT. This component, comprised of two distinct peaks, was maximal over parieto-central areas of the scalp. The double peaked late positivity is apparent in both Congruent and Incongruent conditions. The early initial positivity occurred at about the same time in Congruent and Incongruent conditions. The latency of the second positive peak is about 100 ms longer in the Incongruent condition. The second peak occurred at approximately the same time of the subject's response, which is later in the Incongruent condition than the Congruent condition.

-------- Insert Figure 1 here --------

Discussion

The behavioural results of this study are generally consistent with previous IAT experiments. Although the difference in reaction times between the Congruent and Incongruent conditions was not statistically significant, this result was in all likelihood due to the very small sample size of this study. There was a higher error rate in this study compared to other IAT studies. The original IAT study (Greenwald et al, 1998) found error rates of approximately 0.05. One possible explanation for the high error rate in this study is that the current study presented words relatively rapidly at a fixed interval (ie: every 1200 ms), while the Greenwald paradigm presented each word only after the subject had responded to the previous word. Therefore subjects in the original study had an unlimited amount of time to make their response to each stimulus. This explanation, however, is unlikely since RTs in both studies were quite similar. As indicated in the histograms in Figure 1, subjects rarely needed more than 800 ms in which to make a response. It is also possible that subjects in this pilot study emphasized speed at the cost of accuracy. Faster RTs are, however, typically associated with more errors. In this pilot study, more errors were made in the Incongruent than the Congruent condition, yet RT was still longer in the Incongruent condition. Therefore, the increased error rate cannot account completely for the RT differences. Critically, speed-accuracy response biases do not affect the latency of the P300 component of the ERP (Kutas, McCarthy, & Donchin, 1977). A more likely explanation is that participants in the Greenwald and colleagues (1998) study were given immediate feedback when they made errors. Participants in our study were asked to respond carefully, but were not given immediate feedback on their performance.

The results of pilot study provide support for the use of ERP in studying the IAT effect. The most important finding was the difference in latency and amplitude of the late positive component in the Congruent and Incongruent conditions. This component appears similar to the P300 component of the ERP. The P300 is a positive peak which typically occurs between 250 and 500 ms following stimulus presentation, although the latency can vary greatly depending on the type of task and participant characteristics. The peak observed in this pilot study had a scalp distribution typical of P300, increasing in amplitude from frontal to parietal electrode sites. The positive component observed in the current study had a parietal maximum scalp distribution, and a latency similar to previously reported P300 components. Because a great deal is known about cognitive processes that affect the P300 component of the ERP, this information can inform future studies on the IAT and lead to specific, testable hypotheses.

Conclusion

The pilot study revealed a difference in the latency and amplitude of the P300 between the Congruent and Incongruent conditions of the IAT. There was a decrease in performance in the Incongruent condition compared to the Congruent condition. There also appears to be a decrease in P300 amplitude at Pz in the Incongruent compared to the Congruent condition.

Although the initial pilot study yielded some interesting results that should be examined in subsequent experiments, it is not possible to draw conclusions from this data about the cognitive processes involved in the IAT. This study had a limited sample size of only six participants. More importantly, the number of stimulus presentations ("trials") within specific attribute and target categories was small (50 trials per category for the Congruent and Incongruent conditions). The ongoing EEG includes both the signal (ERP) and background noise activity. EEG averaging can attenuate the background noise allowing the small amplitude ERP to emerge, providing a sufficiently large number of trials are presented. However, only a small number of trials were presented in the pilot study. Background noise was therefore relatively high in individual subjects. Although ERPs are quite clear relative to the background EEG in Figure 1, this represents a grand average – the average of all subject's averages. The ERPs of any individual subject are likely to contain too much error to allow statistical analysis to be completed.

Despite these limitations, the ERP findings of the pilot study provide grounds for testing specific cognitive theories of the IAT. Since both the amplitude and latency of P300 were found to be different in the Congruent and Incongruent conditions, these components will be the focus on subsequent studies.

Chapter 3

Experiment 1

The following article is currently in press in the journal “Neuroreport”, titled: “Event-Related Potential Measures of Processing During an Implicit Association Test”, by Mark A. **Coates** and Kenneth B. **Campbell.**

The Implicit Association Test (IAT) was designed, as its name implies, as an implicit measure of social attitudes (Greenwald et al, 1998). The advantage of the IAT is that while explicit attitudes can be influenced and altered by self-presentation strategies, it is much more difficult to subjectively alter implicit attitudes perhaps occurring outside of conscious awareness. The basic IAT procedure is very similar regardless of the specific type of attitudes being examined. The Implicit Association Test measures strengths of associations between concepts by comparing response times in two combined discrimination tasks. Participants are required to sort stimuli representing four categories (two “target”, two “attribute) using just two responses, each assigned to two of the four categories. The assumption of the IAT is that, if a target and attribute categories are highly associated, the sorting task (as measured by reaction time) will be faster when they share the same response key (Congruent condition) than when they share different response keys (Incongruent condition).

The underlying cognitive processes responsible for the difference in reaction time between the Congruent and Incongruent conditions remain poorly understood (Greenwald, Nosek & Banaji, 2007).. In the present study, event-related potentials (ERPs) were also recorded to provide a real-time measure of the extent of cognitive processing. ERPs consist of a series of negative- and positive-going components, reflecting the extent of information processing. Of particular interest to the study of the IAT is a late parietal maximum, positive component, the P300, so labeled because in very easy perceptual discrimination tasks, its reaches its maximum amplitude at a latency of about 300 ms (Donchin & Coles, 1988). In more complex tasks involving language and/or memory, its latency can be considerably longer (Kok, 1997). Importantly, many studies have now indicated that the latency of the P300 is mainly affected by stimulus classification time and minimally affected by response production time (Donchin & Coles, 1988; Kutas, McCarthy & Donchin, 1977). Several factors affect the amplitude of the P300 (Johnson & Donchin, 1978; Johnson, 1986; Polich, 2007). Almost any rare, task relevant stimulus will elicit a P300. Its amplitude is highly affected by the informative value of the stimulus. If, however, the participant has doubt or equivocation about their decision, the amplitude of the P300 is reduced because the full informative value of the stimulus has not been extracted (Johnson, 1986). P300 is thus typically larger in easy than in difficult tasks. An earlier negative component, N2, is also relevant to an understanding of the processes involved in the IAT. N2 has been interpreted as a reflection of response conflict monitoring (Folstein & Van Petten, 2008). N2 is larger in conditions in which a subject must inhibit a prepotent response, for example, those occurring in Go/NoGo and Stroop tasks (Gajewski, Stoerig & Falkenstein, 2008; Liotti, Woldorff, Perez & Mayberg, 2000) .

Different explanations of the IAT effect have been offered. The original explanation of the IAT (Greenwald, McGhee & Schwartz, 1998) proposed that the Congruent condition would be "easier" to perform than the Incongruent condition. Thus, it might be expected that the amplitude of P300 would be reduced in the Incongruent condition, although precisely what cognitive processes make the Congruent easier remains debated (Greenwald, Nosek & Banaji, 2007). Brendl, Markman, & Messner (2001) argue that in the Incongruent condition of the IAT participants adopt a more stringent decision-making criterion, requiring more extensive evaluation of the stimulus, compared to the Congruent condition. This delay in stimulus classification time would explain the familiar delay in reaction time and should also result in a prolonged P300. Mierke and Klauer (2001) and others (Klauer, Schmitz, Teige-Mocigemba & Voss, 2010) propose that task set switching accounts for the reaction time difference in the IAT. In the Congruent condition, the participant need only categorize both the attribute and target words as being either "good" or "bad" and respond accordingly. In the Incongruent condition, the IAT can be thought of as two separate tasks, one requiring the participant to categorize "good" and "bad" stimuli, and the other requiring categorization of "insect" and "flower" stimuli. On any given trial, participants must first check that the task set used in the previous trial is applicable to the current stimulus. If it is not, they must switch to the other task set in order to be able to correctly respond. Consequently, it would appear that two decisions are made, delaying the response. Complex decisions have been associated with multiple-peaked late positivities (Johnson & Donchin, 1985; Muller-Gass, Gonthier, Desrochers & Campbell, 2000; Girodo, Deck & Campbell, 2000) . In short, the account by Brendl et al (2001) would predict a single but delayed P300 while the Mierke and Klauer task set switching account (2001) would predict two late positivities. In the Incongruent condition of the IAT is the subject needs to inhibit a prepotent response thus delaying reaction time. If this is the case, it would also be expected that N2 would be larger in the Incongruent than in the Congruent condition.

Methods

Participants

Ten healthy young adults (3 men, 7 women) aged 18 to 21 years (M = 18.8years) volunteered to participate in this study. Written informed consent was obtained prior to the study and an honourarium was offered as compensation. This study was conducted following the Canadian Tri-Council (Natural, Health, and Social Sciences) ethical guidelines.

Procedure

Participants were seated approximately 50 cm in front of a computer monitor in a dimly lit, sound attenuated room. They were asked to fixate on a point ("+") in the center of the monitor and to minimize blinking and other eye movements. Participants saw a series of single words presented in black letters on a white background. A total of 20 different words were employed representing one of four categories (Good, Bad, Musical Instruments, and Weapons). Five exemplars were presented within each category. The single exemplar words were selected from a previous IAT study (Mierke & Klauer, 2001) and consisted of 3 to 8 letter words that had no more than 3 syllables. A total of 200 trials was presented in each experimental condition. Order of presentation of the words was randomized. Single words were presented for 300 ms and the inter-stimulus interval (stimulus onset-to-onset) was 1200 ms.

Participants responded to stimuli using either the "Q" or "P" keys on opposite sides of a QWERTY keyboard, with their left and right index fingers respectively. Participants were instructed to keep their fingers on the keys between trials. In the Congruent condition, participants were asked to respond to words they classified as being either "good" or "musical instruments" with the left key and words they classified as being either "bad" or "weapons" with the right key. In the Incongruent condition, participants were told to respond to words they classified as being either "good" or "weapons" with the left key and words they classified as being either "bad" or "musical instruments" with the right key. They were asked to respond as quickly as possible while maintaining a high level of accuracy. No feedback was given for speed or accuracy of responding. Half the participants were presented with the Congruent condition first. Prior to the Congruent and Incongruent conditions, practice conditions were run to familiarize the participants with the responses required in each category. Forty stimuli were presented in each practice run. Upon entry into the testing lab, participants responded to "good" words with the left key and "bad" words with the right key. Prior to the Congruent condition, participants responded to "musical instrument" words with the left key and "weapon" words with the right key. Prior to the Incongruent condition, participants responded to "musical instrument" words with the right key and "weapon" words with the left key. A brief rest period was provided between conditions.

Physiological Recording.

The EEG was recorded from frontal (F3, Fz, F4), central (Cz), parietal (P3, Pz, P4), temporal (T5, T6) and occipital (Oz) scalp sites using silver-silver chloride electrodes attached to an elastic cap. The EEG was referenced to the tip of the nose. A vertical and horizontal electro-oculogram were also recorded. Inter-electrode impedances were below 5 kΩ. Data were acquired continuously at a 256 Hz sampling rate. The high filter was set at 30

Hz and the time constant at 2 s. Offline, an inverse Fast Fourier Transformation high digital filter set at 15 Hz was applied to the data.

Eye movement and blinking artifact were subsequently corrected using an inverse Fourier Frequency Transformation algorithm operating in the frequency domain (Woestenberg, Verbatan & Slangen, 1983). The data were later reconstructed into discrete trials beginning 100 ms before stimulus onset and continuing for 900 ms following it. Only trials in which a correct response was made were averaged.

Data Analyses

Visual inspection of the grand average (average of all participants' averages) waveforms indicated the presence of long latency negative and positive-going deflections, perhaps reflecting N2 and P300 activity. The earlier N2 appeared as two small amplitude peak deflections in the grand average, an early frontal negativity peaking at about 270 ms and a later more posterior negativity peaking at about 340 ms. A scoring method described by Alho et al (1989) was used to quantify these small amplitude waveforms. The post-stimulus epoch was divided into three consecutive 50 ms time windows beginning at 250 ms (i.e., 250-300, 300-350 and 350-400 ms). Within each of these intervals, the average of all data points was computed, yielding an average amplitude measure. The amplitude of the later P300 was much larger and thus measured as the maximum positive peak at Pz (where it was largest) occurring between 400 ms and 600 ms. The amplitude of P300 was measured at other midline sites at a latency corresponding to that of the maximum Pz deflection.

A two-way ANOVA with repeated measures on Condition (Congruent, Incongruent) and Electrode site (Fz, Cz, Pz, Oz) was conducted for the amplitude of the various deflections. A one-way ANOVA with repeated measures on Condition was run on the P300 latency data.

Results

Performance Data

Participants had a higher correct classification rate in the Congruent (M = .86, SD = 0.1) than in Incongruent (M = .74, SD = 0.1) condition, $F(1,9) = 12.30$, $p < 0.01$. Reaction time was also significantly faster in the Congruent (M = 502 ms, SD = 44 ms) than in the Incongruent (M = 575 ms, SD = 58 ms) condition, $F(1, 9) = 14.10$, $p < 0.01$.

Physiological Data.

Grand averages for all channels in the Congruent and Incongruent conditions are illustrated in Figure 2. The amplitude of an initial fronto-central negative wave within the 250-300 ms interval did not significantly vary between Congruent and Incongruent conditions (F<1). Differences first emerged during a later negative deflection in the 350-400 ms interval. A significant interaction was found between condition and electrode site, $F(3, 27) = 3.55$, $p < .05$. This negativity was larger in the Incongruent than the Congruent condition but only at parietal and occipital sites. A large positive component peaking at about 500 ms is visible in both the Congruent and Incongruent conditions. There was no significant difference in the latency of this P300 between the Congruent (M = 486 ms, SD = 30 ms) and Incongruent conditions (M = 495 ms, SD = 32 ms), $F < 1$. P300 was significantly larger in the Congruent than in the Incongruent condition $F(1,9) = 5.54$, $p < .05$. The interaction between Condition and Electrode site was not significant, $F(3,27) = 1.57$, $p > .05$.

-------- Insert Figure 2 here --------

A second later positivity was apparent in the Incongruent condition. This positivity peaked at about 630 ms. Confidence interval testing revealed that this positivity was significantly different ($p < .01$) from the zero voltage pre-stimulus baseline level. This later positive peak was not apparent in the Congruent condition. A 2-way ANOVA was run on the average of all data points between 605 and 655 ms. The mean amplitude within this window did not significantly differ between the Congruent and Incongruent conditions, $F < 1$.

A pre-response lateral readiness potential might also be of interest in this study. The LRP is largest over the scalp contralateral to the hand of movement and is thought to reflect preferential activation/preparation of a specific response. The amplitude of this deflection did not however significantly differ between the Congruent and Incongruent conditions, $F < 1$.

Discussion

The performance data replicate the findings of many previous IAT studies. Reaction time was delayed by about 75 ms in the Incongruent condition.

Differences between the Congruent and Incongruent conditions initially emerged 350-400 ms after stimulus onset, a period corresponding to a late negative N2 peak. The larger amplitude of the N2 in the Incongruent condition is consistent with the interpretation that it provides a measure of response conflict monitoring (Liotti et al, 2000), although there is debate about the specific processes reflected by N2 (Folstein et al, 2008). Moreover, recent research has failed to find N2 differences in a gay-straight version of the IAT, using both words and pictures (Williams & Themansen, 2010). In the present study, the N2 difference was largest over posterior regions of the scalp. It is thus also possible that this N2 difference might be a result of a temporal and spatial summation with the later parietal maximum P300. A large P300 was elicited at about 500 ms by all words. The amplitude of this early positivity was larger in the Congruent condition. It is possible that this overlapping positivity may have “pulled” the earlier N2 in a positive direction, causing it to appear smaller (i.e., less negative-going) in the Congruent condition.

The reduced P300 in the Incongruent condition is in accordance with the hypothesis that this condition is indeed more difficult than the Congruent condition (Greenwald et al, 1998). The attenuation of the early P300 might be attributed to equivocation, a lack of certainty about the initial decision in the Incongruent condition (Johnson, 1986). The latency of the P300 (about 500 ms) corresponded to the time of the behavioural response in the Congruent condition. Importantly, P300 latency did not vary between the Congruent and Incongruent conditions, although reaction time did. Thus, classification of the words in the Incongruent condition does *not* appear to take longer than in the Congruent condition. Similarly, pre-response activation and preparation as measured by the LRP did not differ between the two conditions. What then can account for the reaction time delay in the Incongruent condition? Part of this delay might be attributed to the response conflict monitoring as reflected by N2 differences. Critically, a second small amplitude later positivity was apparent in the Incongruent condition. Because of its small amplitude, this late positivity was difficult to distinguish from noise in the averaged ERP. It is thus was measured as the mean amplitude between 605 and 655 ms. Caution should thus be heeded in interpreting this peak. This double-peaked positivity is nevertheless consistent with the hypothesis that in the Incongruent condition, participants make an initial decision related to the need to either maintain or switch task sets from the previous trial, and subsequently make a second decision about the correct response to the stimulus. Task switching is not required in the Congruent condition and thus only a single, rapid decision needs to be made and therefore a secondary decision is not made. It would appear that the making of this secondary decision on some trials in the Incongruent condition might result in a delayed reaction time. Importantly the average of all data points in the time interval did not significantly differ between the Incongruent and Congruent conditions. In the Congruent condition, the measuring technique might be quantifying a return-to-baseline from the larger and earlier P300. This initial ERP study of the IAT did attempt to closely follow standard procedures and thus a limited number of trials were presented. Unfortunately, a much larger number of trials will need to be presented to allow the possible later positivity to clearly emerge. Further, this would also allow the sorting and averaging of single trials based on whether task switching was required or not.

Conclusions

Several properties of the ERP differed between Congruent and Incongruent conditions of the IAT, including P300 amplitude, N2 amplitude, and the possible presence of a second positive peak in the Incongruent condition. In contrast, P300 latency, a measure of stimulus classification time, did not differ between conditions.

Chapter 4

Experiment 2

Experiment 1 used a musical instrument/weapons version of the IAT. The behavioural results were consistent with those of many previous studies. Accuracy was lower and RTs was longer in the Incongruent than the Congruent condition. ERPs were also collected in an attempt to monitor the cognitive processes that differentiate the Congruent and Incongruent conditions. A late N2, largest at parietal and occipital sites between 350 and 400 ms, was apparent in the Incongruent condition. This might reflect greater response conflict in this condition, although the processes associated with N2 remain much debated (Folstein & Van Petten, 2008). All stimuli in both the Congruent and Incongruent conditions elicited a large amplitude posterior P300, peaking at about 500 ms. The latency to the P300 peak, often claimed to reflect a measure of stimulus classification time, did not however differ between the two conditions. This suggests that the delay in RT in the Incongruent condition must be explained by processes other than stimulus classification time, most likely response-related processes occurring subsequent to the stimulus classification processes. The amplitude of P300 was larger in the Congruent than the Incongruent. This could have been a reflection of equivocation/uncertainty in the Incongruent condition. A second positivity, peaking at about 630 ms was apparent in the Incongruent condition. This second late positivity might also reflect processes that could account for the delay in RT in the Incongruent condition. It was suggested that this second positive peak might reflect the need to make a secondary decision in the Incongruent condition, perhaps related to the frequent need to switch task set strategies (Mierke & Klauer, 2001).

The task set switching account

The task-set switching account is currently considered to be one of the most promising cognitive explanations of the IAT effect. In this account, the difference in RT between the Congruent and Incongruent conditions is the result of task switching costs. The IAT can be conceptualized as two separate tasks, an attribute categorization task and a target categorization task. On any given trial, the participant is required to either categorize an attribute stimulus (for example, Good versus Bad, the attribute categorization task) or categorize a target stimulus (for example, Musical Instrument versus Weapon, the target categorization task), the response keys being shared for the two tasks. Because the response keys are shared between the two tasks, the participant must first make an initial decision about the present stimulus to decide which of the two tasks must be performed, and a subsequent secondary decision to determine which response key to press.

When two categories that are highly associated with each other (for example, “good” and “musical instrument”, or “bad” and “weapon”) are paired on the same key (the Congruent condition), the participants does not actually have to switch between the two tasks because a heuristic shortcut allows the task to be performed correctly. The valence information of the target stimuli (e.g.: "Musical Instrument" = “good”) can be used to determine the correct response. As such, a participant can correctly complete the task without categorizing the targets at all, simply by relying on the valence (good-bad) information contained in any individual stimulus. Rather than determining which of the two tasks needs to be performed, a participant can thus make a single decision based the valence (good or bad) of the stimulus. This single decision might account for the fact that only a single P300 peak was observed in the Congruent condition in Experiment 1. In the Incongruent condition, on the other hand, a participant is not able to rely on a single decision. In this condition, the participant must make two decisions. When a stimulus word is presented (e.g. “tuba”), the participant must first decide whether to categorize it on the basis of its attribute (good-bad) or on the basis of its target (musical instrument-weapon). A secondary decision must then be made. Assuming the participant (correctly) identifies this as a target (as opposed to an attribute), a second decision (“musical instrument”) must be made. As such, additional time is required to switch back and forth between the attribute categorization task and the target categorization task. The authors suggest that the IAT effect is therefore a function of the extra time required to switch between two different tasks in the Incongruent condition. The need to switch between the two tasks results in the double decision and can perhaps account for the double-peaked P300, observed in Experiment 1. A problem with the double-peaked P300 in Experiment 1 is that the peak was very small and might have been attributable in background noise rather than being true ERP component. Moreover, at the time of its occurrence, the amplitude of this second P300 did not differ from the positivity observed in the Congruent condition. The positivity in the Congruent condition might however reflect a return to baseline of the earlier large P300.

The need for task-set switching results in a subtle subtle sequential effect (Mierke & Klauer, 2001). A trial is called a switch trial if it requires a different strategy from that used in the previous trial; it is a repeat trial if the same task and strategy as that used in the previous trial can be used. Although task switching objectively occurs equally often in the Congruent and Incongruent conditions, the participant need not actually switch strategies during the Congruent condition. This is not the case in the Incongruent condition and it is this frequent need to switch task strategies that is hypothesized to account for the delayed RT in this condition Mierke and Klauer (2001, 2003; Klauer & Mierke, 2005). These authors found slower response latencies and more errors for switch than for repeat trials in the Incongruent condition. In the Congruent condition, task set switching costs were minimal. Indeed, IAT effects (longer RT, lowered accuracy) still occur (but are reduced) when task cues precede each trial and allow participants to predict and prepare for the upcoming task, minimizing the need for switching strategies at the onset of the attribute or target stimulus (Mierke & Klauer, 2001).

Mierke and Klauer (2001) identified two types of switching, task-set switching and response switching. On any given trial, a participant could be required to make either a task or a response switch, both a task and a response switch or neither a task nor a response switch (i.e., repeat the same strategy used in the previous trials). For example a task switch would occur if the subject is required to make an attribute decision (happiness = good) on one trial followed by a target decision (trumpet = musical instrument) on the next. A response switch would occur if the subject needed to respond with one key (say the "P") on one trial but a different key ("Q") on the next. A task and a response switch would occur if the subject is presented with a target decision on one trial (rifle = weapon) requiring perhaps a response with the "Q" key and an attribute decision on the next (death = bad) requiring a response perhaps with the "P" key. Mierke and Klauer (2001, 20003) emphasized the importance of task switching trials, which was found to have a larger effect on RT than response switching trials (approximately 140 ms difference in task switch trials; approximately 60 ms difference in response switch trials). They argued that this was due to the additional executive control demands of switching between cognitive sets (or strategies).

However, there was an interaction between response switching and task switching. RT was shortest on trials that required no switch (neither a response switch nor a task switch), and longer for all trials requiring a switch. Therefore, although the authors emphasized the role of executive control processes (task switching), it appeared as though either type of switch resulted in an increase in RT.

A possible double-peaked P300 was observed in Experiment 1. As mentioned, this might reflect the need to effectively switch task strategies more often in the Incongruent condition. The amplitude of the initial P300 was larger in the Congruent condition. The subjective amount of information extracted from the stimulus words might have contributed to the P300 differences observed in the IAT. The ability to rely on a single decision in the Congruent condition may result in less uncertainty (i.e., less equivocation) about a decision. Again, this is because in the Congruent condition, the participants may be able to make use of the valence information ("good" or "bad") from every stimulus, including the target stimuli (weapons = bad; flowers = good) to make their decision. In the Incongruent condition, responding based on the valence information would lead a subject to make the wrong response. Thus, using the valence of a weapon ("bad") may lead to the wrong response, because the target category, weapon, is now paired with the attribute "good". The target category "flower" is now paired with the attribute category "bad". Because only a single decision need be made in the overall congruent condition but a double decision needs to frequently be made in the Incongruent condition, there is greater probability of uncertainty of the decision in the Incongruent condition. Equivocation (uncertainty) results in a reduced P300 (Johnson, 1986).

Sequential relationships between types of stimuli are integral to the task set switching account of the IAT effect (Mierke and Klauer, 2001). Participants respond faster when a stimulus is preceded by another stimulus in the same category (same task) or another stimulus that requires the same behavioural response (same response), and perform slower when the stimulus requires them to either change task set (switch task) or the response key (switch response). This may be due to participant's expectation that subsequent trials will require the same task set as the previous trial or subsequent responses will be the same as the previous response The major factor that affects the amplitude of P300 is the probability of stimulus occurrence (Johnson, 1986), the rarer the occurrence of the relevant stimulus, the larger the amplitude of P300. A classic theory to explain this is that the amplitude of P300 reflects the updating of working memory. The memory for a frequently occurring stimulus is well-established. By comparison, the memory for an infrequently stimulus is poor. Thus, when the rare stimulus is presented, considerably more updating of working memory is required. The extent of updating of working memory is directly related to the amplitude of P300. The overall probability, or what has been called "global" probability, of each category of stimuli is however identical (p = 0.25) in the Congruent and Incongruent conditions. P300 is however also influenced by subtle "local" probability of occurrence (Polich, 2003). Thus in the random sequence a, b, a, a, a, b, b, b, a, b, b, a, the global probability of occurrence of a the a and b stimuli is identical (p = .50). However, within the trials, there is a sequence in which the "a" stimulus occurs on three consecutive trials, followed by a "b". The memory for the "a" improves upon this sequential repetition and thus, the amplitude of P300 would gradually decrease upon repetition of the same stimulus category. On the other hand, immediately following, the 3 consecutive "a" stimuli, the "b" stimulus occurs. Because the memory for it will now be weaker, its occurrence after the string of the "a" stimuli will result in a large P300. Thus, both global and local probabilities within a sequence will affect the amplitude of P300. In the Congruent and Incongruent conditions, task and response switching occurs with equal objective probability, just as overall global probability of attribute and target category presentation is equal. However, in the Mierke and Klauer account, subjects do not actually need switch strategies in the Congruent condition because they can rely on a single task strategy (e.g., happy = good on one trial; daisy = good on the next trial), even though objectively a task switch has occurred. In the Incongruent trials, subjects do need to often switch between tasks (happy = good on one trial; daisy = flower on the next). Because the subject can rely on a single task set strategy in the Congruent condition, local probability of occurrence is less relevant, even though objectively task switching does occur. This is not the case in the Incongruent trials. When the subject encounters an attribute category on one trial and then a target category on the next, they need to switch strategies. This effectively changes local probabilities. Thus, on switch trials (when new information is provided and memory must be updated) , it is possible that the P300 may be larger than on non-switch trials (when information is redundant), but only in the Incongruent trials.

A major limitation of the studies conducted by Mierke and Klauer (2001, 2005) is that only performance data (RT and accuracy) were quantified. RT reflects the overall contribution of several processes (stimulus evaluation, stimulus classification and decision making, response selection and initiation, movement time), but does not allow researchers to examine the contributions of each component process separately. Previous experiments have investigated the ERPs associated task switching, and provide a reasonably strong basis for interpreting the ERPs associated with the IAT, and of the task set switching account proposed by Mierke and Klauer (2001). Target-locked P300 amplitude has typically been found to be smaller on task switching trials than on trials in which a task switch is not required (Goffaux, Phillips, Sinai, & Pushkar, 2006, 2008; Kieffaber & Hetrick, 2005; Perianez & Barcelo, 2009). Notably, many ERP studies (Perianez & Barcelo, 2009; Karayandis, Provost, Brown, Paton & Heathcote, 2010; Barcelo, Escera, Corral & Perianez, 2006; West, Langley & Bailey, 2010) employ a cue stimulus to indicate that a task switch is required, and switch cues have consistently been found to elicit a larger P300 than cues indicating no-switch. During the IAT, the participants do not receive a cue when a task switch is needed. The target itself conveys the need for a task switch. That is, when a target is presented the participant must first use the stimulus to determine whether or not to switch tasks, and then must use that same information to determine the correct behavioural response.

The present study will again use the same musical instrument-weapon version of the IAT. The number of trials will be doubled. This will then permit the sorting and averaging of trials according to whether a task (or response) switch had occurred. Mierke and Klauer (2001, 2003) have previously observed that RT is delayed on task switch trials. RT is also delayed, but to a much lesser extent on response switch trials. In the previous Coates and Campbell (2010) study, N2 was slightly but significantly larger on the overall Incongruent trials, perhaps reflecting response conflict. This effect would be expected to be larger on switch trials within the Incongruent condition. P300 was much larger in the overall Congruent than in the Incongruent condition. Again, subtle effects of the task switching effects may emerge, but only within the Incongruent condition. A small double-peaked P300 was apparent in the Incongruent condition. This might also be related to the need for a switch of task strategy and the need for a double decision in the Incongruent condition

Participants

Participants were 17 university students (10 female). The mean age of participants was 19.5 years (*SD* = 2.0). All had normal or corrected-to-normal vision. Two participants were excluded from the analysis because the data were unusable due to excessive blinking and movement. None of the participants reported using nicotine on a regular basis. None of the participants included in the analysis had consumed more than their usual amount of caffeine or nicotine on the day of the experiment.

All participants provided informed consent before the experiment was conducted and received an honorarium to compensate them for their time. This study was conducted following the guidelines of the Canadian Tri-Council (Medical, Natural Science and Social Sciences) on ethical conduct for research involving humans.

Stimuli

Stimuli consisted of four categories of words (Good, Bad, Musical Instruments, and Weapons), each of which consisted of 5 exemplars (see Appendix 2) used in the previous experiment (Coates & Campbell, 2010) These individual stimuli were selected from a previous IAT study (Greenwald et al, 1998), and were considered to be “common” words based on a representative sampling of English text (Francis & Kucera, 1982). Items were presented in black letters on a white background at the center of a computer monitor at a viewing distance of approximately 50 cm. Stimulus duration was again 300 ms and the inter-stimulus interval (stimulus onset-to-onset) was 1200 ms. The first target acquisition condition, consisted of 25 trials from each of the two target categories (Musical Instruments and Weapons), for a total of 50 trials. Again, there were 5 exemplars in each category and each exemplar was presented 5 different times. Stimuli were presented in random order. The second block “attribute acquisition condition” consisted of 25 trials from the “good” category and 25 trials from the “bad” category, again, a total of 50 trials, again presented in random order. The attribute acquisition condition was followed by one of the two mixed condition (Congruent or Incongruent). Each individual exemplar was now presented 40 times in both the Congruent and Incongruent conditions. Therefore in the congruent in the Congruent condition, 200 trials from each category were presented (i.e., a total of 800 trials was presented). The 800 trials were presented in random order. After the Congruent or Incongruent condition (whichever was presented first), the participants received the target reversal condition, which consisted of 25 trials from each of the two target categories (Musical Instruments and Weapons), for a total of fifty trials. Finally, 800 trials of the other mixed condition (Congruent or Incongruent condition) were presented.

Procedure

A modified IAT paradigm was administered similar to that used in the previous study. During the experiment, participants were asked to fixate on a central fixation point on a computer monitor, thus avoiding eye movement. In all conditions, participants were instructed to respond as quickly as possible while maintaining accuracy. Consistent with previously published studies using the IAT, participants responded to stimuli using one key (“Q”) on the left side of a keyboard and another key (“P”) on the right side of a keyboard. In all conditions, participants used the index finger of the left and right hands respectively to respond. Participants were instructed to keep their fingers on the keys between trials.

In the target acquisition condition, participants were instructed to press the “q” key when the word presented was a musical instrument, and the “p” key when the word presented was a weapon. In the attribute acquisition condition, participants were instructed to press the “q” key when the word presented was a good word, and the “p” key when the word presented was a bad word. In the Congruent condition, participants were instructed to press the “q” key when the word presented was either a musical instrument or a good word, and the “p” key when the word presented was either a weapon or a bad word. In the Incongruent condition, participants were instructed to press the “q” key when the word presented was either a weapon or a good word, and the “p” key when the word presented was either a musical instrument or a bad word. In the target reversal condition, participants were instructed to press the p” key when the word presented was a musical instrument and to press the “q” key when the word was a weapon.

The order of the Congruent and Incongruent conditions was counterbalanced to minimize sequence effects. Half of the participants received the Congruent condition first and half received the Incongruent condition first. This was not expected to make a significant difference, as the procedure followed previously published recommendations by minimizing the effect of condition order by increasing the number of practice trials (Nosek, Greenwald & Banaji 2005)

EEG recording

The EEG was recorded from frontal (F3, Fz, F4), central (Cz), parietal (P3, Pz, P4), temporal (T5, T6) and occipital (Oz) scalp sites using silver-silver chloride electrodes attached to an elastic cap. The EEG was referenced to the tip of the nose. A vertical and horizontal electro-oculogram were also recorded. Inter-electrode impedances were below 5 kΩ. The high filter was set at 30 Hz and the time constant at 2 s Data were acquired continuously at a 256 Hz sampling rate. Offline, the data were subsequently filtered at 15 Hz using an inverse FFT algorithm.

Eye movement and blinking artifact were subsequently corrected also using an inverse FFT algorithm (Woestenberg, Verbatan, & Slangen, 1983). The data were later reconstructed into discrete trials beginning 100 ms before stimulus onset and continuing for 900 ms following it. Only trials in which a correct response was made were averaged.

Trials were sorted and averaged initially on the basis of condition, with all trials in the Congruent condition averaged separately from all trials in the Incongruent condition. Within each condition, trials were subsequently sorted and averaged on the basis of whether or not a particular trial required either a task switch, and then sorted and average again on the basis of whether or not a response switch was required. Notably, response and task switching were independent of one another. For example, a trial requiring a task switch might or might not also require a response switch.   Since Mierke and Klauer (2001) originally found that no-switch trials elicited much shorter RTs than any of the switch trials, it is possible that the effects of task switching and response switching contaminated trials of the other type. Therefore the trials were also sorted into no-switch trials, task switch trials (no response switch), response switch trials (no task switch), and task/response switch trials (both task and response switch).

The amplitude and latency of the positive peak was measured by manual detection of the highest positive peak at Pz between 400 ms and 600 ms in individual subjects. The second peak in the Incongruent condition was not prominent in the grand averages , although the P300 positivity did appear to be extended in the Incongruent condition. An alternate scoring method was used for the scoring of this extended late positivity. A scoring method described by Alho, Sams, Paavilainen, Reinikainen, and Näätänen (1989) is often used to quantify long-lasting waveforms in which a distinctive peak is not apparent. The average of all data points between 600 ms and 700 ms was computed, yielding an mean amplitude measure for each individual subject. Since the P300 peak was largest at Pz, this electrode site was used in all analyses.

N2 amplitude was measured in the 350-400 ms post-stimulus epoch that showed the greatest difference between Congruent and Incongruent conditions in the previous study Again, because a distinctive peak was not always apparent, the Alho et al. (1989) scoring method was used to quantify N2. The average of all data points between 350 and 400 ms was computed, yielding an average amplitude measure. Again, N2 was quantified at Pz, where it was largest.

Statistical Analyses

The performance data (RT and accuracy) were analyzed using a 3-way ANOVA with repeated measures on Condition (2 levels, Congruent and Incongruent), Task set switching (2 levels, switch and no-switch) and Response switching (2 levels, switch and no-switch).Because of the complexity of the ANOVAs (which included additional factors such as task and response switching), amplitude and latency ANOVAs were run only on the Pz data, where the amplitude of N2 and P300 was largest. Because previous experiments (Mierke & Klauer, 2001, Klauer & Mierke, 2005) have shown that set switching processes differ in the Incongruent but not within the Congruent condition, P300 amplitude was further analyzed using planned comparisons to examine the differences between all switching trials in the Incongruent and no-switch trials of the Incongruent.

Results

Behavioural results

A significant difference in both the speed of response and accuracy and was again found between the overall Congruent and Incongruent conditions. RT was significantly longer in the Incongruent condition $F(1,14) = 53.87$, $p < 0.001$ (M =580 ms, SD = 66 ms) than in the Congruent condition (M = 512 ms, SD = 61 ms). Similarly participants made significantly more errors in the Incongruent condition (M = 0.22 errors, SD = 0.06) than in the Congruent condition (M = 0.14, SD = 0.04), $F(1,14) = 15.36$, $p < 0.002$.

Performance was then examined within each condition. A two-way interaction between condition and task set switching was found for RT, $F(1,14)= 13.9$, $p <0.003$. Task switching had no effect in the Congruent condition (for task switch: M = 517 ms, SD = 90 ms; for no task switch: M = 509 ms, SD = 100 ms). Task switching did however significantly prolong RT in the Incongruent condition (for task switch: M = 593 ms, SD = 102 ms; for no task switch: M = 567 ms, SD = 112 ms). The main effect of response switching was not significant, $F(1,14) = 1.84$, $p > 0.24$. A significant interaction between task set switching and response switching was however found, $F(1,14) = 18.04$, $p <0.001$. RT was different on trials that did not require a response switch depending on whether a task switch was required (M = 555 ms, SD = 105 ms) or also not required (M = 528 ms, SD = 88 ms), while on trials that required a response switch there was no difference between trials that also required a task switch (M = 551 ms, SD = 103 ms) and those requiring no task switch (M = 549 ms, SD = 99 ms). The 2-way interaction between response switching and condition was not significant $F<1$. The 3-way interaction was not significant. To further explore the differences between trials types noted by Mierke and Klauer (2001) within the Incongruent condition, trials that contained any type of switch were compared to trials that did not contain a switch. RT was significantly longer for switch trials (M = 582.41 ms, SD = 71.25) than no-switch trials (M = 557.40, SD = 63.72), $t(14) = 4.81$, $p < 0.001$. For all analyses, similar results were obtained using median RT.

Similar results were observed for error rate. The main effect of task set switching was significant, $F(1,14)= 24.55$, $p < 0.001$, while response set switching was not significant, $F<1$. Again, task switching also interacted with other effects. A significant two-way interaction between condition and task set switching, $F(1,14)= 8.21$, $p <0.02$ was again found. In the Congruent condition, error rates between task switching (M = 0.15, SD =0.11) and no switching trials (M = 0.13, SD = 0.13) did not differ. However, in the Incongruent condition, trials containing a task switch resulted in a higher error rate (M = 0.24, SD = 0.18) than in trials that did not require a task switch (M = 0.18, SD = 0.16). Again, a significant interaction between task set switching and response switching was found, $F(1,14) = 32.43$, $p <0.0001$. The proportion of errors made by participants was different on trials that did not require a response switch depending on whether a task switch was required (M = 0.13, SD = 0.11) or not required (M = 0.22, SD = 0.15), while on trials that also required a response switch there was no difference between trials requiring a task switch (M = 0.17, SD = 0.14) and those requiring no task switch (M = 0.18 errors, SD = 0.13). The 2-way interaction between response switching and condition was not significant, $F<1$. The three way interaction was not significant, $F<1$. Within the Incongruent condition alone, trials that required a switch were associated with a significantly higher error rate (*M* = 0.33, *SD* = 0.11) than no-switch trials (*M* = 0.11, *SD* = 0.09), $t(14) = 4.94$, $p < 0.001$.

ERP results

The overall grand averages observed in the Congruent and Incongruent conditions are presented in Figure 3. These results were consistent with the results observed in Experiment 1. The overall ERP consisted of early (0-300 ms) negative and positive components, probably related to sensory processing of the visual stimuli. These were followed by a later N2, peaking between 350 and 400 ms that was maximum over parietal regions. A large parieto-central P300 was elicited at about 550 ms. This positivity was extended considerably in the Incongruent condition, although this did not appear as a distinctive secondary peak.

-------- Insert Figure about 3 here --------

Figure 4 isolates only ERPs elicited in the Congruent condition, but trials are sorted and averaged according to whether or not an objective task switch occurred, and also whether or not a response switch occurred. Very little difference can be seen between the different trial types in the Congruent condition. Figure 5 isolates ERPs elicited in the Incongruent condition and again are sorted on the same basis as those in Figure 4. N2 differences are not visible in between any trial types in the Incongruent condition. P300 amplitude appears somewhat larger in the task switch and response switch trials than on trials without each type of switch.. On all trial types in the Incongruent condition the peak of P300 is longer lasting than in the Congruent condition, although a distinct double peak is not evident in any trial type.

-------- Insert Figures 4 and 5 about here --------

To further explore the effects of task and response switching, trials of the Incongruent condition in which any type of switch occurred (task, response, or task/response) were averaged together and compared to those trials that did not contain any switching in Figure 6. The P300 peak is notably smaller on trials that contained no switch compared to trials that required any type of switch.

-------- Insert Figure 6 about here --------

The main effect of condition on the N2 amplitude in the 350-400 ms interval was not significant although there was a trend $F(1,14) = 2.74$, $p < 0.12$ trend towards larger negativity in the Incongruent (for amplitude: $M = 1.06$ μV, $SD = 0.75$ μV) than the Congruent condition (for amplitude: $M = 1.81$ μV, $SD = 0.75$ μV), The 2-way interactions between condition and task switching, condition and response switching, and task and response switching were not significant (all $F < 1$). A complex 3-way interaction was however significant $F(1,14) = 5.41$, $p < 0.04$. The results of this interaction were however complex and were not predicted on the basis of the literature.

A distinctive early P300 was elicited in both the Congruent and Incongruent conditions. P300 latency was significantly longer in the Incongruent condition ($M = 541$ ms, $SD = 24$ ms) than in the Congruent condition ($M = 503$ ms, $SD = 22$ ms), $F(1,14) = 6.29$, $p < 0.05$. There were no other significant main effects or interactions involving P300 latency.

A significant main effect of condition was found for the amplitude of P300, $F(1,14) = 14.3$, $p < 0.003$, Amplitude was reduced in the Incongruent condition ($M = 4.40$ μV, $SD = 2.35$ μV) compared to the Congruent condition ($M = 5.86$, $SD = 2.21$). A significant two-way interaction was however found between condition and task set switching, $F(1,14) = 7.78$, $p < 0.03$. No other two-way or three-way interactions were significant.

Within the Incongruent condition, trials which required a switch (either a task switch, response switch, or both) were compared to trials that did not require a switch. This planned comparison was undertaken to further explore the account of Mierke and Klauer (2001, 2003), who found that the shortest RTs occurred on trials that did not require any switching, and that all types of switching were associated with longer RT. While the comparison failed to reach significance, there was a very strong trend towards larger amplitude P300 on switch trials (M = 9.23 μV, SD = 5.18) than on no-switch trials (M = 7.78 μV, SD = 5.15), $t(14) = 1.67$, $p < 0.06$.

A distinct second P300 peak was not visible in the Incongruent condition. Within the Incongruent condition, confidence interval testing was used to determine if the later, extended positivity significantly differed from the zero voltage pre-stimulus baseline voltage. Confidence interval testing determines the probability that the minimum value of this component reach a non-positive (negative) value (i.e., the minimum value is less than the 0 μV baseline). The late positivity was significantly different from the zero voltage pre-stimulus baseline level on all trials within the Incongruent condition regardless of whether a task or a response switch occurred or if no switch occurred. A 2-way ANOVA was conducted to explore the effects of task switching and response switching on this peak. The effect of task switching was not significant, $F(1, 14)=2.54$, $p < 0.15$, while response switching was significant $F(1, 14) = 5.71$, $p < 0.05$. The interaction between task and response switching was not significant. Post-hoc analysis revealed that trials requiring no switch (M = 2.06 μV, SD = 2.01 μV) differed significant from trials requiring a task switch alone (M = 2.91 μV, SD = 1.82 μV), $t(14)= 2.57$, $p <0.02$, a response switch alone (M = 3.30 μV, SD = 1.95 μV), $t(14)= 2.47$, $p < 0.02$, and trials that required both a task and response switch (M = 3.18 μV,

SD = 2.31 μV), $t(14)$= 2.43, $p < 0.02$. Amplitude at this interval did not differ significantly between any of the types of switching trials.

Discussion

Behavioural Data:

The behavioural results of the current study were consistent with many past studies of the IAT. Thus, error rates were higher and RT was delayed by about 70 ms in the overall Incongruent condition compared to the Congruent condition. This is very similar to the results observed in the previous experiment (Coates & Campbell, 2010) run under almost identical conditions, except for the addition of additional trials in the present study. Importantly, switching did not have an effect on RT or accuracy in the Congruent condition. This was not the case in the Incongruent condition. Error rates were higher and RT was longer on switch trials in the Incongruent condition. This also replicates the results of other studies (Mierke & Klauer, 2001, Klauer & Mierke, 2005), a finding that is used to support the argument that task switching costs are higher in the Incongruent condition than in the Congruent condition.

Physiological Data:

The present study also recorded ERPs to attempt to explain the differences in processing between the switch and no switch trials in the Incongruent condition. The morphology of the overall ERPs in the Congruent and Incongruent conditions were quite similar to those observed in the previous experiment.

N2: A small posterior N2 peaking between 350-400 ms was somewhat larger in the Incongruent condition. This difference was not significant but did follow the same trend observed in the previous study. The Incongruent condition requires participants to avoid the prepotent tendency to respond with the same key for categories that share common characteristics (e.g., both "gun" and "death" would be considered "bad", but the participant must not respond to "gun" as if it were a "bad" word; these target categories were mapped to the "good" key). This likely increased the need for response conflict monitoring in the Incongruent condition. According to the task switching account, this is not required in the Congruent condition because participants can simply respond to all trials using a single task set (categorizing all trials as either "good" or "bad"). A complex 3-way interaction between condition, task switching, and response switching was found. This 3-way interaction is however difficult to interpret because similar 3-way interactions for performance data have not been reported.

P300: RT was significantly longer in the overall Congruent than in the Incongruent condition. The latency of P300 was also longer in the Incongruent condition. This finding differed from the results of the previous study, which did not find a significant difference in P300 latency between conditions. This might be explained by the fact that in the previous study, two distinctive late positive peaks were found in the Incongruent condition. In the present experiment, a long-lasting positivity was apparent in the Incongruent condition, but distinctive peaks were not observed. From a methodological point of view, this makes the scoring of the maximum P300 deflection difficult to identify, resulting in latency jitter. The identification of the timing of the "true" peak of the P300 might thus be somewhat arbitrary. Nevertheless, in the grand averages, it is apparent that the peak of the Congruent P300 does considerably overlap with the initiation of the P300 in the Incongruent condition. RT differences were also found within the Incongruent condition, RT being delayed on task switch trials. This was not the case with P300 latency. No differences in its latency were found across types of trials, even between task switch and no task switch trials in the Incongruent condition. This finding suggests that the delay in RT between switch trials and no-switch trials in the Incongruent condition are more likely a result of processes related to response selection rather than stimulus classification time. Such a response strategy would be consistent with the classification of switch and no switch trials being made at the same time. It is processes that are responsible for the subsequent selection, initiation and execution of the response that results in the delay of the RT in the task switch trials. Doubt (equivocation) about the confidence of the decision and a subsequent delay (perhaps double-checking) in making the response could thus result in an overall delayed RT. It would appear that the making of this secondary response-related decision (perhaps a double-check before execution of the response, or perhaps a secondary decision related to the actual response to be made on some trials in the Incongruent condition might result in a delayed reaction time.

Overall, a larger amplitude P300 was observed in the Congruent condition than in the Incongruent condition. This might be because of uncertainty (equivocation) related to doubt about the response decision in the Incongruent condition. The less demanding nature of the Congruent condition could result in greater working memory resources being available to the process the stimuli, resulting in greater certainty about the correct response to the stimulus. Goffaux et al (2006) found that a smaller amplitude P300 was elicited during conditions that required task set switching than in conditions that did not. Since P300 amplitude is believed to increase proportionally as target processing is facilitated, a process affected by the amount of available working memory resources (Johnson, 1984; Kok, 2001), a larger target-locked positivity on repeat compared to switch trials was thought to indicate the availability of a larger amount of working memory resources that could be used to process repeat targets. Similarly, if greater working memory resources are required in the Incongruent condition than in the Congruent condition of the IAT, available stimulus processing resources will be greater in the Congruent condition and will result in less uncertainty and a larger amplitude P300.

A within condition analysis indicates subtle differences between the Congruent and Incongruent conditions. The Mierke and Klauer (2001) account suggests that within the Congruent trials, processing strategy does not differ between task switching and non-task switching trials and between response switching and non-response switching trials. The ERP data are in agreement with this hypothesis. In the Congruent condition, P300 amplitude was not different between task switching and non-task switching trials or between response switching and non-response switching trials. As Mierke and Klauer suggest, subjects might therefore have relied on a single stimulus classification strategy that probably relies simply on an attribute categorization (good or bad) for both actual attribute (happy = good; corpse = bad) and target (rifle=bad; rose = good) stimuli. Thus, all trials in the Congruent condition may well have been processed using an identical strategy and classification scheme even though objectively, participants were asked to use different classification strategies for the attribute and target categories. In the Incongruent condition, on the other hand, the results were opposite of predictions made based on performance data. The task switching account suggests that the longer RTs on the task switch trials can be explained by an increase in task difficulty and caused by the need to now switch decision strategies. It was predicted that the largest amplitude P300 would occur on the trials in which no switch was required, and the smallest P300 should occur on trials requiring a task set switch. In reality, all switching trials (response switch and task switch) elicited an equally large amplitude P300, while "no-switch" trials elicited a significantly smaller amplitude P300. Because P300 amplitude within the Incongruent condition was not consistent with the findings of previous studies of task set switching, this implies that other factors are involved in the IAT effect. The results of the current experiment can be interpreted in the context of the triarchic model of P300 amplitude, which describes the factors that can affect P300.

The major factor affecting the amplitude of P300 is the probability of occurrence of the task-relevant stimulus. A stimulus that occurs often is well-represented in working memory. On the other hand, a stimulus that occurs rarely is poorly represented. When it is presented, the working memory for this stimulus must be updated. It is this updating of working memory that can be reflected in the amplitude of P300 (Donchin & Coles, 1988). The two types of attribute stimuli and two types of target stimuli occur with equal probability (p = 0.25) in both the Congruent and Incongruent conditions. Therefore, a priori probability of individual stimuli could not have been a factor in this experiment, since all stimuli were presented an exactly equal number of times. The probability of switch and no-switch trials could have conceivably affected P300 amplitude, since there was a slight difference in the number of trials of each type. If this were the case, the least probable types of trial (no switch = 46 trials; task switch, no response switch = 45 trials) would have elicited the largest amplitude P300, while the most frequently occurring types of trials (response switch, no task switch = 54 trials; combined response/task switch = 55 trials) would have elicited the smallest P300. The results were nearly opposite of what this argument would predict; the no-switch trials of the Incongruent condition elicited the smallest amplitude P300, while all switching trials elicited larger P300s. Therefore it can be concluded that a priori probability likely did not have an effect on the P300 amplitude differences observed.

The effect of sequential expectancy, or local probability, can readily explain the amplitude differences between different types of trials within the Incongruent condition. In the Incongruent condition, a larger amplitude P300 occurred when a change in task set is required. Switching to a new task set on that trial would not be consistent with the sequential expectancy created by the previous trials. That is, a participant responding to several trials in a row that all require the use of the same task set would assume that the local probability of this trial type is high, and this assumption would be violated by the first trial to require a task switch. This change in local probability would result in a larger amplitude P300. In the Congruent condition all trials can be completed with the same task set. Therefore it is unlikely that there would be any effect of sequential expectancy within this condition.

While local probability can account for the amplitude differences within the Incongruent condition, it does not offer insight into why the P300 elicited on all trials of the Congruent condition is significantly larger than any of the trials of the Incongruent condition. As described above, however, overall differences between the two conditions can be explained by greater uncertainty (equivocation) in the Incongruent condition. Therefore, the effect of sequential expectancy can explain the amplitude differences across trial types within the Incongruent condition, while uncertainty can explain the overall difference between Congruent and Incongruent conditions.

Interestingly, there was no difference found between task set and response set switching trials in the Incongruent condition. This finding runs counter to the predictions of the task set switching account, which suggests that task set switching relies more heavily on executive control processes than response switching trials, and thus should elicit different

ERPs. However, if we assume that more uncertainty occurs throughout the Incongruent condition rather than at the level of individual stimuli, ERP results within the Incongruent condition can be accounted for by local probability. All switching trials, whether task set switching or response switching, represent a change in local probability. That is, a participant responding with the left hand key for several trials in a row would expect to respond with the same key on the next trial as well. If the subsequent stimuli then required the right hand key to be pressed, this would represent a change in local probability.

The results of the current experiment, while at first glance appearing to contradict the task set switching account in some respects, are fairly consistent with this model. Increased response conflict monitoring, as reflected in N2 amplitude, was evident in the Incongruent condition. A difference in the level of uncertainty in the two conditions is a likely explanation for the overall difference in P300 amplitude between conditions. In the Congruent condition all trials are likely processed in a similar manner, as the task set switching account would suggest. In the Incongruent condition the effect of task set switching is reflected in smaller P300 amplitude compared to the Congruent condition. Within the Incongruent condition, differences in local probability across trial types appear to account for the lower P300 amplitude in trials that do not require a switch. An extended P300 was noted in the Incongruent condition, possibly representing a double decision on all trials of this condition.

Conclusion

The current results appear to suggest that the interaction between condition and trial type is the result of the combined effects of uncertainty and local probability. The greater degree of equivocation in the Incongruent condition, explains the amplitude difference between the Congruent and Incongruent conditions. The effect of sequential expectancy explains the amplitude differences between trial types in the Incongruent condition. These factors affect P300 amplitude to a different degree; equivocation has a multiplicative effect on P300 amplitude, while sequential expectancy has an additive effect. It is therefore not surprising that the Congruent condition elicits a much larger P300 than the switch trials of the Incongruent, since the difference in equivocation has a larger effect on amplitude than the difference in sequential expectancy.

What implications does the current study have for the task set switching account of the IAT effect? The task set switching account implies that the no-switch trials of the Incongruent condition are processed in a similar fashion as all trials of the Congruent condition, since the participant is not required to switch task set on either of these types of trials. This was clearly not the case, as different ERPs were elicited in the Congruent condition from those elicited by the Incongruent no-switch trials. Instead, it appears that differences in processing across conditions are due to the great uncertainty in all trials of the Incongruent compared to all trials of the Congruent. The results of the current experiment imply that the difference in use of executive control processes occurs at the level of each condition, as well as within different trial types of the Incongruent condition.

Although the preceding discussion is consistent with the ERP results, future experiments will need to be conducted to directly test this theory. The effects of local probability depend on differences in working memory usage across trial types of the Incongruent condition. One way of testing whether working memory processes differ across trial types would be to present irrelevant stimuli which are expected to cause distraction while participants perform the IAT. This would draw attention from the visual IAT task, and presumably interfere with the ability of participants to utilize the resource limited working memory system. This is the basis for the final experiment in this thesis.

Chapter 5

Experiment 3

Working memory and the task set switching account

In Experiments 1 and 2, small differences in N2 amplitude were observed between the Congruent and Incongruent conditions, possibly reflecting greater response conflict monitoring in the Incongruent condition. Overall, a larger amplitude P300 was observed in the Congruent condition than in the Incongruent condition. Experiment 2 examined the task set switching account of the IAT effect by sorting and averaging the trials of IAT task on the basis of whether a task or response switch had occurred, and analyzing the ERPs associated with each type of trial. This allowed the ERPs associated with task set or response switching trials to be distinguished from those associated with no-switch trials. P300 was largest in the Congruent condition, smaller on task set or response switching trials of the Incongruent condition, and smallest on trials of the Incongruent condition that did not require a switch.

P300 has been the subject of many studies and reviews. Donchin (1981) in a classic review, developed the notion that P300 was intimately related to the updating of working memory. P300 is usually studied in the so-called "oddball task" in which a subject is presented with a frequently occurring stimulus and at odd (or rare) and unpredictable times, a "target" that is to be detected. The P300 is only elicited if the subject actively attends to the stimulus train and detects the target stimulus. This is because the memory for the frequently presented stimulus is thought to be stored in working memory and is maintained by active rehearsal. By comparison, the memory for the rarely presented target is poor. Thus, when the target is presented, the memory for it must be updated. It is the extent of updating of memory that determines the amplitude of P300. Thus, the amplitude of P300 increases inversely with the probability of target occurrence. However, as Johnson (1986) noted, amplitude of P300 is also influenced by factors other than objective probability, including local probability and degree of uncertainty (equivocation).

The previous experiment noted subtle differences between switch and no switch trials, but only in the Incongruent condition. In the Congruent conditions, trials that required a switch of tasks or response modes elicited essentially identical P300s as those that did not require a switch. This was not the case in the Incongruent condition; within the Incongruent condition, switching trials elicited similarly large amplitude P300s, while trials that did not require a switch elicited significantly smaller amplitude P300s. This difference in P300 amplitude between switch and no-switch trials in the Incongruent condition was explained by local probability. The subject's strategy for making a decision did not need to change on no-switch trials. On the other hand, the strategy did need to change if either a task or a response switch was required. Obviously, to know if a switch was warranted required a memory of strategies used in the past. Presumably, this information was stored in working memory and this working memory needed to be updated when either a task or a response switch was required. Therefore, the results of the previous studies suggested that the updating of working memory, as reflected in the amplitude of P300, might have been responsible for differences between switch and no-switch trials. Uncertainty about the correct response to make as reflected in overall P300 amplitude (and perhaps response conflict monitoring as reflected in the amplitude of N2), differed between the Congruent and Incongruent conditions.

Previous studies provide strong support for the notion that task set switching and working memory are both critically important in the IAT (Klauer et al, 2009). Differences in the need to update working memory were suggested as a possible explanation for the P300 variance between switch and no switch trials in Experiment 2. Experiment 3 will test further this hypothesis by manipulating working memory. Manipulating working memory while participants are actually engaged in the IAT could be a fruitful way to determine the relative contributions of task set switching and working memory. To that end, the final experiment will alter working memory by introducing an auditory distractor to the IAT procedure.

Distraction and working memory

The IAT requires active attention to be sustained for the entire duration of the task. Previous studies have indicated that the switching of the task or the response strategies will have an important effect on performance in the Incongruent condition, and a minimal effect in the Congruent condition (Mierke & Klauer, 2001). Information hypothetically retained in working memory about the strategy used on previous trials will thus have an important interactive effect. Working memory is sustained by active rehearsal (Baddeley, 2003; Lavie & de Fockert, 2005). Cognitive tasks require resources, but these resources are limited. A "central executive" is thought to be responsible for the allocation and control of these resources. A number of studies have now indicated that certain highly relevant, but otherwise unattended stimuli can interrupt the functioning of the central executive (Berti, Roeber, & Schroger, 2004; Munka & Berti, 2006; SanMiguel, Corral & Escera, 2008). Such input forces the central executive to cease (or alter) the processing of current ongoing cognitive tasks and switch attention to the processing of the interrupting input. The processing of the potentially irrelevant input diverts resources away from ongoing cognitive tasks, leading to a deterioration in performance. This is called “distraction”. The purpose of this “attention capture” mechanism is to allow the observer to become aware of critical input, even if attention is not active directed toward it. In nature, this allows for the detection of biologically relevant information, such as stimuli indicating the presence of predators or prey.

The auditory modality is particularly effective at interrupting the central executive. This is because we can hear sounds that occur outside of the focus of vision (i.e., beside and behind our heads). Naatanen has developed and updated an extensive and elaborate model of the auditory “experience” (Naatanen, Kujala, & Winkler, 2010). This model proposes two modes by which attention capture is possible in the auditory modality. A transient detector system, as its name implies, detects transient aspects of the stimulus, such as its onset or offset. The output of this system varies directly with the transient energy (typically “intensity”). Naatanen has employed ERPs as a means to measure the extent of processing of unattended input. The amplitude of a negative component, N1, is claimed to reflect the output of the transient detector system. The amplitude of N1 varies directly with the intensity of the auditory stimulus. If the output of the transient detector system reaches a certain critical level, operations of the central executive are interrupted and attention is switched to the auditory channel.

The switching of attention is thought to be reflected by another component of the ERP, P3a (Escera et al., 1998). P3a differs from the P300 in that is elicited passively (unlike the P300 which requires active attention), has a centro-frontal scalp topography (unlike the parietal maximum of P300) and occurs 250-300 ms after stimulus onset (unlike the 300-600 ms of P300). Because it occurs after P3a, P300 has also come to be known as "P3b".

The second mode of attention capture involves the operations of the change detector system. It detects the change in any feature of an auditory stimulus including its frequency, duration, location or intensity.

The oddball paradigm is often used to study the operations of the change detection system. In it, the subject is presented with a frequently occurring, "standard" stimulus. At odd and unpredictable times, a feature of the standard is changed to form a "deviant". The features of the standard are stored in a brief-lasting sensory memory. The features of an incoming stimulus are compared to those that exist as a representation in sensory memory. If they match (as is the case when the standard is presented), the strength of the memory store improves. When the deviant is presented, its features fail to match those of the sensory representation and change is detected. The output of the changed detector system is measured by a different negativity, the Mismatch Negativity (MMN), peaking between 100 and 200ms. This output varies directly with the extent of stimulus change. If the output is sufficiently large, functioning of the central executive will again be interrupted, resulting in the switch of attention as indicated by the P3a.

A deviant that signals an increase in intensity from the standard is particularly effective at disrupting the central executive. This is because the deviant activates the transient detector system more than the standard (the deviant is louder); in addition, the deviant also activates the change detector system (it signals change). Thus, the central executive may be interrupted by signals arising from either or both of the two different routes.

Rinne, Sarkka, Degerman, Schroger & Alho (2006) presented participants with an auditory duration discrimination task. Participants were asked to discriminate between tones of different frequencies. At rare times, the intensity of the stimulus was either increased by as much as 9 dB (an increment) or decreased by as much as 9 dB (a decrement). The irrelevant increment resulted in much larger distraction than the decrement, as evidenced by a larger deterioration in performance on the duration-detection task. Also, the increment elicited a large amplitude late positivity, P3a, peaking at about 250-300 ms. The decrement did not. Others have noted that a 20 dB decrement from an 80 dB standard will elicit a P3a but this is still much smaller than that elicited by a 10 dB increment (Muller-Gass et al., 2007, Macdonald, Jamshidi, & Campbell, 2008).The effectiveness of intensity increments in capturing the attention may be because an increase in stimulus intensity is experienced as "looming" or approaching and may elicit a fear response in animals, while a decrease in intensity is experienced as "fading" or retreating (Bach, Schachinger, Neuhoff, Esposito, Di Salle, Lehmann, et al., 2007; Freiberg, Tually, & Crassini, 2001; Maier & Ghazanfar 2007). So effective is an intensity increment that it can even elicit a P3a in sleeping participants (Macdonald et al., 2008).

A highly relevant stimulus input might thus result in a "bottom-up" forceful switch of attention. One of the functions of the central executive is thought to be protection from distracting stimuli (Postle, 2005). That is, the central executive functions as a "top-down" modulator of information received from sensory systems (Naatanen et al, 2010). Auditory distractibility has been shown to be modulated by the amount of cognitive load of a concurrent attentional task (Lavie, Hirst, De Fockert, & Viding, 2004; Lavie, 2010). There is however much debate about the interaction of the cognitive demands of an attended task and the effects of distraction. Lavie and de Fockert (2005) indicated that distraction occurs most often for cognitive tasks that require the updating of working memory and less often for tasks that simply involve a perceptual decision about a current stimulus (i.e., there is no need for the updating of working memory). On the other hand, some studies have indicated that when a cognitive task makes greater demands for central executive resources the effects of distraction by an irrelevant stimulus are decreased (Escera et al., 2000; Berti & Schroger, 2003; SanMiguel, Corral, & Escera, 2008). Presumably, fewer resources are available for the processing of the irrelevant stimulus. Other studies have, however, found no effect of even very heavy workload demands on effects of interruption of the central executive as measured by the amplitude of P3a (Muller-Gass et al., 2006; 2007).

The effects of distraction under different working memory loads has implications for the current research into the IAT effect. Mierke and Klauer (2001, 2004) first described the different costs associated with executive processes during different conditions of the IAT. These authors have shown that switch trials were associated with performance declines (increased RT and error rate) compared to no-switch trials, which they believed was due to the need for additional engagement of central executive resources. The purpose of the present study is to determine if the IAT effect can be modulated by unattended, but potentially disruptive increases in the intensity of a frequently occurring auditory stimulus. In Experiment 2, it was postulated that this might also be related to differential needs for the updating of working memory. According to Lavie (2010), a task that requires information to be held in working memory will be more susceptible to distraction than a task that does not. An irrelevant auditory deviant should interfere with the updating process more than a frequently presented standard. The previous study suggested that updating of working memory is less important on no-switch trials than on switch trials in the Incongruent condition, and less important in the Congruent condition than in the Incongruent condition.

On each trial of the Incongruent condition, participants must decide whether or not to switch strategies based on information contained within the stimulus itself (ie: whether the category of the stimulus in current trial is the same as or different than the previous trial). Most ERPs studies that employ auditory stimuli as distractors present them either before or concurrently with the task relevant stimuli. This would not be appropriate in the case of the IAT. This is because in the IAT, the decision to switch strategies can only occur *after* presentation of the visual stimulus and a decision is made as to the nature of the task. RTs on switch trials are about 600 ms after stimulus onset.. Thus, for auditory distractors to interfere with the decision to switch strategies, processing of the auditory stimulus should be concurrent with the decision about the nature of the task (i.e., between 400 and 700 ms after the onset of the visual stimulus). ERP components such as the N1 and MMN are though to reflect automatic, obligatory (pre-attentive or pre-conscious) processing of auditory stimuli, and occur within the first 100-200 ms after stimulus onset. An optional P3a to an increment deviant will be be elicited at about 250-300 ms, if it results in a switch of attention. (Naatanen, 1990). Therefore, in the present study, the auditory stimuli were presented 300 ms after the onset of the visual stimuli. The P3a to the increment would thus occur at about 550-600 ms after the onset of the task relevant visual word

Methods

Participants

Participants were fifteen university students (nine female). The mean age of participants was 22.7 years (*SD* = 2.6 years). None reported a history of neurological or psychiatric disorders. All reported normal or corrected-to normal-vision. Participants were asked to abstain from alcohol and caffeine use for 24 h prior to the start of the study. Written informed consent was obtained prior to the study and an honourarium was offered as compensation. This study was conducted following the Canadian Tri-Council (Natural, Health, and Social Sciences) ethical guidelines. The EEG data from one participant were excluded because of a large amount of artifact as a result of frequent blinking and movement. The data from a second participant were excluded due to an inability to perform the task correctly. This participant's proportion of correct responses approached chance level in all conditions.

Stimuli

Stimuli were the same as those used in the previous studies and consisted of four categories of words (Good, Bad, Musical Instruments, and Weapons), each of which consisted of 5 exemplars (see Appendix 2). These individual stimuli were selected from a previous IAT study (Greenwald et al, 1998), and were considered to be “common” words based on a representative sampling of English text (Francis & Kucera, 1982). Items were presented in black letters on a white background at the center of a computer monitor at a viewing distance of approximately 50 cm. Stimulus duration was again 300 ms and the inter-stimulus interval (stimulus onset-to-onset) was 1200 ms. The first target acquisition condition, consisted of 25 trials from each of the two target categories (Musical Instruments and Weapons), for a total of 50 trials. Since there were 5 exemplars in each category, there were 5 presentations of each individual stimulus item. Stimuli were presented in random order. The second block "attribute acquisition condition" consisted of 25 trials from the "good" category and 25 trials from the "bad" category, again, a total of 50 trials, presented in random order. The attribute acquisition condition was followed by one of the two mixed condition (Congruent or Incongruent). Each individual exemplar was presented 40 times in both the Congruent and Incongruent conditions. Therefore, in the Congruent condition, 200 trials from each category were presented (i.e., a total of 800 trials was presented). The 800 trials were presented in random order. After the Congruent or Incongruent condition (whichever was presented first), the participants were administered the target reversal condition, which consisted of 25 trials from each of the two target categories (Musical Instruments and Weapons), for a total of fifty trials. Finally, the other mixed condition (Congruent or Incongruent condition) was presented.

Auditory stimuli were presented binaurally through ER-3A insert earphones. They were presented 300 ms after the onset of the word stimuli. They had a frequency of 1000 Hz and a duration of 50 ms and a rise-and-fall time of 5 ms. The auditory stimuli were irrelevant to the visual IAT task. Importantly, on 85% of trials, the intensity of a frequently presented tone was 70 dB SPL. On the remaining 15% of trials, the intensity of the "standard" was changed to a loud, obtrusive 90 dB SPL "deviant". The standard and deviant tones were presented in a pseudo-random order, under the constraint that a deviant tone could not follow another deviant tone. Previous studies have indicated that a rare 10 dB increase in intensity from a standard level will elicit both behavioural (deterioration in performance) and physiological evidence (P3a) of distraction (Rinne et al.; 2006; Muller-Gass et al., 2007).

Procedure

A modified IAT paradigm was administered which was identical to that used in the previous study. During the experiment, participants were asked to fixate on a central fixation point on a computer monitor, thus avoiding eye movement and minimize eye blinking. In all conditions, participants were instructed to respond as quickly as possible while maintaining accuracy. Consistent with previous studies using the IAT, participants responded to stimuli using one key (“Q”) on the left side of a keyboard and another key (“P”) on the right side of a keyboard. In all conditions, participants used the index finger of the left and right hands respectively to respond. Participants were instructed to keep their fingers on the keys between trials.

In the target acquisition condition, participants were instructed to press the “Q” key when the word presented was a musical instrument, and the “P” key when the word presented was a weapon. In the attribute acquisition condition, participants were instructed to press the “Q” key when the word presented was a good word, and the “P” key when the word presented was a bad word. In the Congruent condition, participants were instructed to press the “Q” key when the word presented was either a musical instrument or a good word, and the “P” key when the word presented was either a weapon or a bad word. In the Incongruent condition, participants were instructed to press the “Q” key when the word presented was either a weapon or a good word, and the “P” key when the word presented was either a musical instrument or a bad word. In the target reversal condition, participants were instructed to press the “P” key when the word presented was a musical instrument and to press the “Q” key when the word was a weapon.

The order of the Congruent and Incongruent conditions was counterbalanced to minimize sequence effects. Half of the participants were administered the Congruent condition first and half were administered the Incongruent condition first. This was not expected to make a significant difference, as the procedure followed previously published guidelines (Nosek, Greenwald & Banaji 2005) on minimizing the effect of condition order by increasing the number of practice trials.

Physiological recording

The EEG was recorded from frontal (F3, Fz, F4), central (Cz), parietal (P3, Pz, P4), temporal (T5, T6) and occipital (Oz) scalp sites using silver-silver chloride electrodes attached to an elastic cap. The EEG was referenced to the tip of the nose. Vertical and horizontal electro-oculogram were also recorded. Inter-electrode impedances were below 5 kΩ. The high filter was set at 30 Hz and the time constant at 2 s. Data were acquired continuously at a 256 Hz sampling rate. Offline, the data were subsequently filtered at 15 Hz using an inverse FFT algorithm.

Eye movement and blinking artifact were subsequently also corrected using an inverse FFT algorithm (Woestenberg, Verbatan, & Slangen, 1983). The data were later reconstructed into discrete trials beginning 100 ms before onset of the visual word and continuing for 900 ms following it. Only trials in which a correct response was made were averaged.

All trials were averaged to the onset of the visual stimulus. Trials were sorted and averaged first on the basis of condition, with all trials in the Congruent condition averaged separately from all trials in the Incongruent condition. Within each condition, trials were subsequently sorted and averaged into “switch trials” and “no-switch trials” on the basis of whether or not the trials required the participant to make a switch (either a task switch, response switch, or both). All switch trials were averaged together, since the previous experiment found evidence of similar processing in all switch trials. Trials were also sorted and averaged on the basis of whether a deviant or standard tone preceded the visual stimulus.

A problem with presenting the standard and deviant auditory stimuli after the IAT stimuli is that the averaging epoch would include the ERPs to both the visual and auditory stimuli. Indeed the ERP components relevant to this study, the N2, peaking at about 350 ms and the P300, peaking at about 500-600 ms will summate both temporally and spatially with those to the auditory stimuli. N2 amplitude was measured in the post-stimulus epoch which showed the greatest difference between Congruent and Incongruent conditions in Experiment 1 (350-400 ms). Within this interval, a scoring method described by Alho, Sams, Paavilainen, Reinikainen, & Näätänen (1989) was used to quantify these small amplitude waveforms. The average of all data points was computed, yielding an average amplitude measure. Since N2 differences were most prominent at Pz in the previous study, this electrode site was used in all analyses.

Visual inspection of the grand average waveforms was used to determined the latency of the late positive peak following presentation of the task relevant word stimuli. The amplitude and latency of the positive peak were measured by manual detection of the highest positive peak at Pz between 450 ms and 600 ms. This peak was analyzed at Pz, the electrode site at which it was the most prominent in the previous experiments. The late positive peak was also analyzed at Cz, since the P3a related to auditory distraction has a more centro-frontal scalp distribution than the P3b.

Statistical analyses

Two separate 3-way repeated measures ANOVAs were run on the dependent variables RT and error rate. Independent variables were condition (2 levels, Congruent and Incongruent conditions), trial type (2 levels, "switch trials" requiring a task and/or response switch, and "no-switch" trials requiring neither type of switch), and auditory stimulus (2 levels, deviant and standard). Based on the results of the previous study, no-switch trials in the Incongruent condition were expected to result in shorter RT and lower error rate than switch trials. Thus, two separate one-tailed t-tests were conducted for the dependent variables RT and error rate to investigate the effects of switching within the Incongruent condition. Similarly, trials containing a standard auditory stimulus in the Incongruent condition were expected to result in shorter RT and lower error rate than trials containing a deviant stimulus. Two separate one-tailed t-tests were conducted for the dependent variables RT and error rate to investigate the effects of standard and deviant tones within the Incongruent condition. It was expected that RT would be longer and error rate higher on trials that contained a deviant tone.

The ERPs to the task relevant IAT word stimuli did overlap and summate with those to the standard and deviant auditory stimuli. Because the standards and deviants were physically different, the resulting visual-auditory composite ERP will vary depending on which specific auditory stimulus was presented. Thus, ERPs in which an auditory standard was presented could not be directed compared to those in which an auditory deviant was presented. These ERPs might vary because of the sensory processing of the auditory stimuli rather than the psychological-cognitive processing associated with the IAT. On the other hand, comparing only averages in which the standard was elicited assures that its (the standard's) influence on the composite ERP will be constant. Similarly, the effects of the deviant stimulus on the composite ERP should be constant in all Congruent and Incongruent conditions. To analyze the effects of condition and switching on trials containing identical auditory stimuli, separate 2-way repeated measures ANOVAs for standard and deviant auditory stimuli were run on four dependent variables: N2 amplitude, P300 latency, and P300 amplitude at Cz and Pz. Independent variables were condition (2 levels, Congruent and Incongruent conditions), and trial type (2 levels, switch and no-switch). Finally, separate 2-way repeated measures ANOVAs for the Incongruent and Congruent condition were run on the dependent variables N2 amplitude, P300 latency, and P300 amplitude at Cz and Pz. Independent variables were trial type (2 levels, switch and no-switch), and auditory stimuli (2 levels, deviant and standard).

Results

Behavioural results

There was a significant performance difference between the Congruent and Incongruent conditions. RT was significantly longer in the Incongruent condition ($M = 665$ ms, $SD = 93$ ms) than in the Congruent condition ($M = 580$ ms, $SD = 74$ ms), $F(1,12) = 16.46$, $p < 0.001$. The main effect of task set switching was significant, $F(1,12) = 27.83$, $p < 0.001$, with longer RT on switch trials ($M = 644$ ms, $SD = 91$ ms) than on no-switch trials ($M = 604$ ms, $SD = 103$ ms). The main effect of the auditory stimulus was not significant $F < 1$.

There was a significant 2-way interaction between condition and switching, $F(1,12) = 19.95$, $p < 0.01$. Switch trials resulted in longer RT than no-switch trials in both the Congruent and Incongruent conditions, but this difference was greater in the Incongruent condition (for switch: $M = 695$ ms, $SD = 109$ ms; for no-switch: $M = 638$ ms, $SD = 124$ ms) than in the Congruent condition (for switch: $M = 592$ ms, $SD = 73$ ms; for no-switch: $M = 570$ ms, $SD = 83$ ms). No other 2-way or 3-way interactions were significant. Within the Congruent condition alone, RT was not significantly longer on trials containing deviant tones ($M = 584$ ms, $SD = 82$ ms) compared to those containing standard tones ($M = 579$ ms, $SD = 74$ ms), $t(12) = 1.13$, $p = 0.14$. Similarly, within the Incongruent condition, RT was also not significantly longer on trials containing deviant tones ($M = 666$ ms, $SD = 120$ ms) compared to those containing standard tones ($M = 667$ ms, $SD = 114$ ms), $t < 1$.

Similar results were observed for error rate. Participants made significantly more errors in the Incongruent condition ($M = 0.18$, $SD = .08$) than in the Congruent condition ($M = 0.09$, $SD = .06$), $F(1,12) = 4.98$, $p < 0.05$. The main effect of task set switching was significant, $F(1,12) = 6.33$, $p < 0.05$, with more errors being made on switching trials ($M = 0.20$, $SD = 0.12$) compared to non-switching trials ($M = 0.16$, $SD = 0.02$). The main effect of auditory stimuli was not significant. None of the interactions were significant. Within the Congruent condition alone, error rate was not significantly higher on trials containing deviant tones ($M = 0.13$, $SD = 0.08$) compared to those containing standard tones ($M = 0.15$, $SD = 0.07$), $t(12)=1.62$, $p<0.10$. Within the Incongruent condition, error rate was similar on trials containing a deviant auditory stimulus ($M = 0.22$, $SD = 0.11$) and on trials containing a standard auditory stimulus ($M = 0.21$, $SD = 0.11$), $t(12) < 1$.

Physiological results

The overall ERPs elicited in the Congruent and Incongruent conditions (Figure 7) appeared generally similar to those obtained in the previous experiment. Differences are however apparent because of the overlapping and summating influences of the ERPs elicited by the auditory stimuli 300 ms after the onset of the IAT word stimuli. ERPs following presentation of the word stimuli again consisted of a series of early negative and positive components, probably related to sensory processing of the visual stimuli. These were followed by a later N2, peaking between 350 and 400 ms that was maximum over parietal regions and slightly more negative-going in the Incongruent condition. A large parieto-central P300 was elicited at about 550 ms in both conditions, with larger amplitude visible in the Congruent condition than the Incongruent condition. A double peaked P300 was visible at posterior sites in both the Congruent and Incongruent conditions, occurring at approximately 550 ms and 650 ms in both conditions. This was not apparent in other studies, but it is difficult to determine the extent to which this was caused by the processing of the auditory stimuli. In the Incongruent condition, a slower return to baseline following the P300 peak was visible.

--------Insert Figure 7 here-------

Within the Congruent and Incongruent conditions, trials containing a standard auditory stimulus were subtracted from trials containing a deviant auditory stimulus in order to isolate the ERPs elicited only by the standard and deviant auditory tones. Figure 8 illustrates the difference wave computed by subtracting the “raw” standard ERPs from the "raw" deviant ERPs. The subtraction computation removes processing that is common to both the standards and deviants (i.e. processing of the IAT word stimuli), leaving only the differential processing of the auditory deviant. As may be observed, a large negativity was apparent at about 450 ms after the onset of the IAT words (i.e., 150 ms after the onset of the auditory stimuli) in both the Congruent and Incongruent conditions. This negativity represents the additional processing of the auditory deviant and probably reflects a composite N1 and MMN. However, a later positivity, the P3a was not apparent at any electrode site in either the Congruent or Incongruent conditions.

--------Insert Figure 8 here-------

Trials were also sorted and averaged on the basis of the type of auditory stimulus that followed the visual IAT stimulus (standard or deviant), condition (Congruent or Incongruent, and type of trial (task switch or no-switch). Figure 9 isolates midline ERPs elicited in the Congruent condition when standard and deviant tones were presented. Trials are sorted and averaged according to whether or not an objective switch occurred. For trials in which a standard auditory stimulus was presented, there appears to be slightly greater negativity of N2 on no-switch trials than switch trials. Very little difference in P300 amplitude is apparent between switch and no-switch trials. For trials in which a deviant auditory stimulus was presented, no-switch trials elicit a larger amplitude P300 than switch trials at frontal and central sites. For both standard and deviant trials of the Congruent condition, the double peaked P300 at Pz is visible for both switch and no-switch trials.

--------Insert Figure 9 here-------

Figure 10 isolates ERPs elicited in the Incongruent condition when standard and deviant tones were presented. Trials are sorted and averaged according to whether or not an objective switch occurred. Similar to trials in the Congruent condition on which standard tones were presented, on trials of the Incongruent containing a standard tone there appears to be a slightly greater negativity of N2 for no-switch trials compared to switch trials. P300 amplitude was also somewhat smaller on no-switch trials. A double peaked positivity was visible at Pz for both switch and no-switch trials. When deviant tones were presented, differences in N2 amplitude were not apparent, possibly due to a relatively high level of background noise, because the deviant stimulus was presented infrequently. There was also little visible difference in P300 amplitude between switch and no-switch trials.

--------Insert Figure 10 here-------

N2: There were no significant main effects of either condition, switching or auditory stimulus for N2 amplitude. Similarly, none of the interactions of these variables was significant. However, the interaction between switching and auditory stimulus approached significance $F(1,12) = 3.56$, $p < 0.08$. For trials in which the auditory deviant was presented, N2 amplitude was larger on switch trials ($M = 1.70$ µV, $SD = 4.00$ µV) than on no-switch trials ($M = 0.62$ µV, $SD = 6.2$ µV). In contrast, for trials in which the auditory standard was presented, N2 amplitude was smaller on switch trials ($M = 0.51$ µV, $SD = 3.34$ µV) than on no-switch trials ($M = 1.64$ µV, $SD = 3.54$ µV).

P300: The difference in latency of P300 in the Congruent (M = 613 ms, SD = 25 ms) and the Incongruent (M = 611 ms, SD = 24 ms) conditions was not significant, F<1. There was however a significant main effect of type of auditory stimulus on P300 latency, $F(1,12) = 23.76$, $p < 0.01$. The presentation of deviant tones (M = 620 ms, SD = 25 ms) resulted in longer P300 latency than the presentation of standard tones (M = 604 ms, SD = 24 ms). The 2-way interaction between auditory stimulus and condition was also significant, $F(1,12) = 14.49$, $p < 0.01$. In the Congruent condition, deviant tones were associated with slightly longer latency P300 (M = 616 ms, SD = 26 ms) than standard tones (M = 610 ms, SD = 17 ms), whereas in the Incongruent condition, the differences between latency on trials containing standard (M = 599 ms, SD = 16 ms) and deviant tones (M = 624 ms, SD = 27 ms) was much larger. None of the other interactions were significant.

The results of the 3-way ANOVA indicated that the main effect of condition on P300 amplitude approached significance, $F(1,12) = 4.50$, $p < 0.06$. P300 amplitude was larger in the Congruent condition (M = 6.68 µV, SD = 0.72 µV) than in the Incongruent condition (M = 5.28 µV, SD = 0.83 µV). There was a significant main effect of switching $F(1,12) = 8.66$, $p < 0.02$, with the larger amplitude on switch trials (M = 6.76 µV, SD = 0.75 µV) compared to no-switch trials (M = 5.19 µV, SD = 0.75 µV). The main effect of auditory distraction was not significant $F(1,12) < 1$. P300s did not differ on trials in which an auditory deviant was presented compared to when a standard was presented. None of the 2-way interactions were significant, nor was the 3-way interaction significant ($F < 1$ for all interactions).

Analyses of amplitude of P300 at Pz were also conducted separately for deviant and standard tones, using 2-way ANOVAs. For trials containing standard tones, the main effect of condition was not significant $F < 1$. The amplitude of P300 did not significantly differ between Congruent and Incongruent conditions. There was however a significant main effect of switching $F(1,12) = 5.24$, $p < 0.05$. Across both Congruent and Incongruent conditions, switch trials elicited a larger amplitude P300 ($M = 5.35$ μV, $SD = 3.29$ μV than no-switch trials ($M = 4.97$ μV, $SD = 3.62$ μV). The interaction between condition and switching was not significant, $F < 1$. For trials containing deviant tones, the main effects of condition $F(1,12) = 2.16$, $p < 0.16$, and switching $F(1,12) = 1.40$, $p < 0.29$, were not significant. The interaction between condition and switching was not significant, $F < 1$.

P300 amplitude at Cz was also analyzed separately for deviant and standard tones. For trials containing standard tones, the main effect of condition was not significant $F < 1$. There was trend towards an effect of switching $F(1,12) = 3.39$, $p < 0.1$. The interaction between condition and switching was not significant. For trials containing deviant tones, no main effects or interactions were significant, $F < 1$ for all effects.

Discussion

Behavioural Data:

The behavioural results of the current study were consistent with many past studies of the IAT. Thus, error rates were higher and RT was delayed by about 70 ms in the overall Incongruent condition compared to the Congruent condition. This is very similar to the results observed in the previous two experiments run under almost identical conditions, except for the addition of the auditory stimuli in the current experiment. Switching had an effect on RT and accuracy in both the Congruent and the Incongruent conditions, but this effect was larger in the Incongruent condition. This finding replicates the results of

Experiment 2 and previous research (Mierke & Klauer, 2001; Klauer & Mierke, 2005), although in the current study the effect of switching was greater in the Congruent condition than in previous experiments. Surprisingly, the effect of auditory stimuli was not significant in either the Congruent or Incongruent condition. Previous studies have indicated that 10 dB increases in intensity from a auditory standard do result in significant deterioration in performance (Rinne et al., 2006). The nature of their task was however quite different from the present one. Their subject were engaged in an auditory duration discrimination task and at rare times, the intensity of the auditory stimulus was increased. Thus, the auditory stimulus itself was task relevant although the increase in intensity was not. In the present study, subjects were engaged in a visual IAT task but the irrelevant auditory stimulus was presented well after (300 ms) after the word stimulus. It is possible that the deviant auditory stimuli were not sufficiently different from the standards to cause an involuntary shift in attention away from the visual task. Within the Congruent condition, there was a very slight trend towards slower performance on trials containing deviant auditory stimuli, which might indicate that distraction caused by the auditory stimuli were too small to be detected in a study with such a small sample size.

Physiological data:

The morphology of the overall ERPs in the Congruent and Incongruent conditions were quite similar to those observed in the previous experiment. A small posterior N2 peaking between 350-400 ms was slightly larger in the Incongruent condition. This difference was not significant but did follow the same trend observed in the previous study. A P300 peak was larger in the Congruent condition than in the Incongruent condition, similar to the results of previous studies.

Subtracting ERPs elicited in trials in which a standard auditory stimulus was presented from those in which a deviant auditory stimulus was presented revealed the presence of large amplitude fronto-central maximum negativity peaking at about 150 ms. This is probably a composite N1 (reflecting greater activation of the transient detector system by the increment deviant) and a MMN (reflecting the fact that the deviant signals change from the past). Importantly, Naatanen insists that both N1 and the MMN are obligatory, automatic responses of the auditory system. The data in the present study are consistent with this. It was expected that the increment would also elicit a large P3a. This was however not observed. Therefore the deviant was not more likely than the standard to cause a shift of attention, or to disrupt the central executive. In light of previous research showing that deviant auditory stimuli can distract from a visual task, this result was somewhat surprising. The P3a is thought to reflect a switch of attention from the current cognitive task and toward the auditory channel (Escera et al., 1998). In the Naatanen model, this interruption of the central executive and subsequent switching of attention is nevertheless considered to be "optional" and can be modulated by a number of factors, including task demands of the primary cognitive task (the IAT in this case). It is possible that because the auditory stimuli were presented after the visual words of the IAT, processing demands of the IAT may have been so extensive the "optional" attention capture by the incidental auditory increment was prevented. Thus deviants did not divert resources away from the IAT more than the standards. It is however important to note that only differential effects of the processing of the auditory deviants could be isolated through the subtraction procedure. It was not possible to determine the actual processing of the standards and deviants separately. Although unlikely, it is possible that standard and deviant auditory stimuli both caused similar levels of distraction and elicited P3a components with similar amplitude. Alternatively, neither the standard not the deviant elicited a P3a. In either case, there was no evidence of greater distraction following presentation of the deviant.

The purpose of the current experiment was to determine if distraction caused by the deviant would differentially affect switch and non-switch trials in the Incongruent condition of the IAT. The task set switching account of the IAT effect would predict that switch and no-switch trials of the Incongruent condition should have been differentially affected by distraction caused by deviant auditory stimuli, since the need for central executive resources was hypothesized to be greater on switch than on no-switch trials. Results indicated no effect of the deviant on performance between switch and no switch trials and little effect on ERPs, contrary to the hypotheses. It was presumed that switch trials would involve more updating of working memory than on no-switch trials in the Incongruent condition. This may not be the case. Lavie (2010) notes that distracting stimuli only have an effect on tasks that require the updating of working memory. It is possible that the IAT makes minimal demands on working memory and thus irrelevant deviant stimuli will have a minimal effect on performance. Alternatively, the introduction of auditory stimuli could have increased the difficulty of the task and caused participants to check and update working memory on all trials of the Incongruent, including no-switch trials on which they could have used the same strategy as the previous trial.

Differences in P300 amplitude between switch and no-switch trials of the Congruent condition were observed. These differences were nevertheless much smaller than in the Incongruent condition. In Experiment 2, there were no significant behavioural or physiological differences between switch and no-switch trials within the Congruent condition. Interpreting these inconsistent findings is therefore difficult. The current results do suggest that different processing takes place on switch and no-switch trials of the Congruent condition, at least when auditory stimuli are presented during the IAT. However, one could speculate that the introduction of the auditory stimuli resulted in verification of strategy in the Congruent condition. The auditory stimuli may therefore have affected participant's ability to use a single strategy in the Congruent condition. Thus, in the Congruent condition participants may have utilized a somewhat similar approach to the task as in the Incongruent condition, which would have resulted in the observed differences between switch and no-switch trials in both conditions. If the auditory stimuli did cause a difference in participants' approach to the Congruent condition, it did not have a notable effect on the overall difference in RT or error rate between the two conditions. This may be because the Congruent condition can be performed relatively automatically, and therefore performance is not extensively impaired even in the context of distraction.

The results of Experiment 3 indicate that the introduction of an obtrusive and irrelevant auditory increment deviant has little overall effect on the IAT, and a similar effect on switch and no-switch trials within the Incongruent condition. There was good evidence that the standard and deviant stimuli were processed differently (N1/MMN), but there was little evidence that the deviant resulted in attention capture. Performance did not deteriorate on the IAT. Further, there was little evidence that the deviant elicited a P3a, a component thought to reflect the actual switching of attention. This could be because both the Congruent and Incongruent conditions of the IAT make such extensive demands on central processing resources that few are available to allow for the switching of attention, or it is possible that the IAT does not require significant updating of working memory. As such, there should be little effect of distraction.

Chapter 6

Summary and Conclusions

The purpose of the current thesis was to use ERPs to investigate the task set switching account of the IAT (Mierke and Klauer, 2001, 2005). Since this research was initiated, there has been increasing interest in the use of ERPs to study the IAT. Indeed, there have been two recently published articles from other laboratories using similar techniques (Hurtado et al, 2009; Williams & Themansen, 2010). However, the current thesis is the first research to test a specific cognitive theory of the IAT effect using ERPs.

The pilot study was a qualitative, exploratory study to determine which components of the ERP might differ between Congruent and Incongruent conditions. A series of ERP components were elicited as participants performed the IAT. A series of small deflections peaking before 300 ms were apparent, but these probably reflected sensory processing of the visual word stimuli. A negative component, N2, peaking at about 350 ms did appear to reveal small differences between the Congruent and Incongruent conditions. A later positivity, peaking at about 550 ms, the P300, was much more prominent. Emphasis was therefore placed on the N2 and P300 in later studies.

The first experiment was an initial follow-up to the pilot using a larger sample. The performance data replicated the findings of many other IAT studies. Thus, RT was longer and error rate was higher in the Incongruent than in the Congruent condition. The study also examined ERPs associated with difference in processing between the two conditions. A larger N2 was found in the Incongruent condition, possibly reflecting greater response conflict monitoring. The latency of P300, thought to reflect mainly stimulus classification processes and to be minimally affected by response-related processes did not differ between the two conditions. Differences in RT between the two conditions might thus be associated with more with response-related than stimulus classification processes. Importantly, significant differences in P300 amplitude were noted. It was significantly attenuated in the Incongruent condition. This was interpreted as a reflection of possible greater equivocation in the Incongruent condition. A second P300 peak, noted only in the Incongruent condition, was interpreted to reflect a double decision. The two decisions in the Incongruent condition might involve deciding which of the two tasks are needed, and then a decision about the categorization of the stimulus. In the Congruent condition only a single decision is needed, since participants do not switch between tasks and need only categorize any stimulus as either "good" or "bad." This interpretation was consistent with the task set switching account of the IAT effect,

The second experiment therefore tested the task set switching account in more detail by presenting a much larger number of trials. This afforded the separation and sorting of trials on the basis of both task and response switching. This was the first ERP study to do so. In the Congruent condition, the amplitude of P300 did not significantly differ across the various trial types. This finding supported the task set switching account, which argues that all trials of the Congruent condition are processed in a similar fashion. This was not the case in the Incongruent condition. Trials requiring any switch (including a task switch, a response switch, or both types of switching) were found to elicit significantly larger amplitude P300 from trials that did not require any switch. Surprisingly, the no-switch trials of the Incongruent condition elicited a smaller P300 than the switch trials. The task set switching account might have predicted that no-switch trials would use less resources (and result in less equivocation) and would therefore elicit a P300 with an amplitude similar to those observed in the Congruent condition. Other factors however influence the amplitude of P300. These findings were interpreted in the context of the triarchic model of P300 amplitude. Local probability, and the consequent checking and updating of working memory, was thought to be responsible for the differences within the Incongruent condition, while increased equivocation throughout the entire Incongruent condition was thought to be responsible for the relatively lower amplitude P300 component compared to all trials of the Congruent condition.

The final experiment therefore examined the role of the presumed differences in the updating of working memory in the different conditions of the IAT. Updating of working memory requires active rehearsal. An irrelevant stimulus can interfere with this updating process. Thus, the third experiment presented two auditory stimuli after the presentation of the IAT word stimuli. A frequently occurring standard should have had minimally interfered with the processing of the IAT. On the other hand, a rare, but obtrusive deviant (an increase in intensity) should have interfered with the updating process. More specifically, the task switching account predicted that deviant tones should have caused greater interference on switch trials than on no-switch trials in the Incongruent condition. The results of Experiment 3 indicated that the introduction of an obtrusive and irrelevant auditory increment deviant did not however have an overall effect on the IAT. Similarly, the deviant did not have a greater effect on switch and no-switch trials within the Incongruent condition. There was good evidence that the standard and deviant stimuli were processed differently (larger N1/MMN following presentation of the deviant), but there was little evidence that the deviant resulted in attention capture (a larger P3a). This could have been because both the Congruent and

Incongruent conditions of the IAT make such extensive demands on central processing resources that few are available to allow for the switching of attention, or it is possible that the IAT does not require significant updating of working memory.

The usefulness of ERPs in the study of the IAT effect is demonstrated by the current research. In particular, the finding that behavioural results were not always consistent with the ERP results demonstrates that electrophysiological measures can complement traditional behavioural measures. The dissociation between behavioural and physiological results was most notable in Experiment 2, in which the P300 amplitude on no-switch trials in the Incongruent condition was significantly different from amplitude on all trials of the Congruent condition, despite RTs on no-switch trials of the Incongruent being shorter than on switch trials and thus more similar to trials of the Congruent condition. This suggests that cognitive operations that may be too subtle to detect using behavioural measures alone may be accessible through the use of electrophysiological measures. For performance differences to emerge using only performance measures often requires the testing of several participants (a minimum of 40-50). The present studies indicate that differences in ERPs can emerge after the testing of a much smaller number of participants (12-15). Of course, ERPs are far more invasive and technically demanding than the behavioural measures.

Original findings

This thesis makes significant contributions to the understanding of the IAT effect. Original findings based on the results of the experiments in this thesis include the following.

1. In all experiments performance data replicated the findings of previous IAT studies,

despite minor methodological changes to the standard IAT procedure and with a much smaller number of participants.

2. In the pilot study, qualitative differences were observed between the ERPs elicited by trials of the Congruent condition and those elicited in the Incongruent condition. Slight differences were noted in amplitude of the P300. A second peak of the P300 component was found to be slightly delayed in the Incongruent condition compared to the Congruent condition.
3. In the first experiment, the amplitude of the late negative N2 peak was found to differ significantly between the Congruent and Incongruent conditions, suggesting that response conflict monitoring may play a role in the IAT effect.
4. Latency of P300 did not differ significantly between the Congruent and Incongruent conditions, suggesting that classification of the words in the Incongruent condition did not take longer than in the Congruent condition. Similarly, pre-response preparation as measured by the LRP did not differ between conditions.
5. Differences in the P300 amplitude between the Congruent and Incongruent conditions were found to be statistically significant. This finding was in accordance with the hypothesis that the Incongruent condition is more difficult than the Congruent condition. The attenuation of the P300 might be attributed to lack of certainty in the Incongruent condition. A double-peaked P300 was observed in the Incongruent condition but not in the Congruent condition, raising the possibility that a secondary decision in the Incongruent condition delays RT.
6. In the second experiment, P300 latency was significantly longer in the Incongruent condition than in the Congruent condition. The positive peak in the Incongruent

condition was less distinct and returned to baseline more slowly than in the Congruent condition, which may explain why latency appeared longer in the Incongruent condition. A distinct secondary positive peak was not however observed in the Incongruent condition.

7. As in the previous experiment, the amplitude of P300 was larger in the Congruent condition than in the Incongruent condition. The less demanding nature of the Congruent condition may have resulted in greater resources being available to process the stimuli, resulting in less uncertainty about the correct response to the stimulus.
8. Contrary to predictions, trials requiring neither a task switch nor a response switch elicited a small amplitude P300, while trials requiring any type of switch elicited a larger amplitude P300. This finding was interpreted as a reflection of greater perceived local probability on no-switch trials compared to switch trials.
9. In the third experiment, participants were presented with irrelevant tones while completing the IAT. Deviant tones were associated with slightly longer P300 latency, although the effect of tones on RT was not significant.
10. Within the Incongruent condition, switch and no-switch trials were similarly affected by deviant and standard tones. This finding is inconsistent with the task set switching account, which predicts a difference in distraction effects on switch and no-switch trials. It is possible that updating of working memory was similar in the Congruent and Incongruent conditions when auditory stimuli were presented. Alternatively, the IAT may not make extensive demands on working memory.

Conclusion

Since its creation in 1998, the Implicit Association Test has become a commonly used measure in social psychology and related fields of research. Studies of the cognitive processes involved in the IAT are necessary to establish the validity of this measure and to suggest further refinements to its use and interpretation. The use of ERPs in the study of the IAT is still in its infancy, but recent research suggests that this is a fruitful method of investigating the IAT effect.

References

Adolphs, R. (2009). The social brain; neural basis of social knowledge. *Annual Review of Psychology,* 60, 693-716.

Aidman, E.V. & Carroll, S.M. (2003). Implicit individual differences: Relationships between implicit self-esteem, gender identity, and gender attitudes. *European Journal of Personality,* 17, 19-36.

Alho K, Sams M, Paavilainen P, Reinikainen K, & Näätänen R. (1989). Event-related brain potentials reflecting processing of relevant and irrelevant stimuli during selective listening. *Psychophysiology,* 26, 514-528.

Asendorpf, J.B., Banse, R,m & Mucke, D. (2002). Double dissociation between implicit and explicit personality self-concept: the case of shy behaviour. *Journal of Personality and Social Psychology*, 83, 380-393.

Awh, E., Vogel, E.K., & Oh, S.-H. (2006). Interactions between attention and working memory. *Neuroscience*, 139, 201-208.

Bach, D., Schachinger, H., Neuhoff, J., Esposito, F., Di Salle, F., Lehmann, C., Herdener, M., Scheffler, K., & Seifritz, E. (2007) Rising sound intensity as an intrinsic warning cue. *Cerebral Cortex*, *18,* 145-150.

Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences*, 4, 417-423.

Baddeley, A. (2002). Working memory: looking back and looking forward. *Nature Reviews: Neuroscience*, 4, 829-839.

Barcelo, F., Escera, C., Corral, M.J., & Perianez, J.A. (2006). Task switching and novelty processing activate a common neural network for cognitive control. *Journal of Cognitive Neuroscience*, 18, 1734-1748.

Barcelo, F., Munoz-Cespedes, J.M., Pozo, M.A., & Rubia, F.J. (2000). Attentional set shifting modulates the target P3b Response in the Wisconsin card sorting test . *Neuropsychologia,* 38, 1342-1355.

Barcelo, F., Perianez, J.A. & Knight, R.T. (2002). Think differently: a brain orienting response to task novelty. *Cognitive Neuroscience and Neuropsychology*, 2002, 13, 1887-1892.

Barnes-Holmes, D., Staunton, C., Barnes-Holmes, Y., Whelan, R., Stewart, I., Commins, S., Walsh, D., Smeets, P.M., & Dymond, S. (2004). Interfacing Relational Frame Theory with cognitive neuroscience; semantic priming, the Implicit Association Test, and event related potentials. *International Journal of Psychology and Psychological Therapy,* 4, 215-240.

Barnes-Holmes, D., Murtagh L., Barnes-Holmes, Y., & Stewart, I. (2006). Using the Implicit Association Test and the implicit assessment procedure to measure attitudes towards meat and vegetables in vegetarians and meat-eaters. *The Psychological Record,* 60, 287-306.

Bartholow, B.D., Riordan, M.A., Saults, J.S., & Lust, S.A. (2009). Psychophysiological evidence of response conflict and strategic control of responses in affective priming. *Journal of Experimental Social Psychology,* 45, 655-666.

Berti, S., & Schroeder, E. (2003). Working memory controls involuntary attention switching: evidence from an auditory distraction paradigm. *European Journal of Neuroscience,* 17, 1119-1122.

Berti, S., Roeber, U., & Schroeder, E. (2004). Bottom-up influences on working memory: behavioural and electrophysiological distraction varies with distractor strength. *Experimental Psychology*, 51, 249-257.

Brendl, C.M., Markman, A.B., & Messner, C. (2001). How do indirect measures of evaluation work? Evaluating the inference of prejudice in the Implicit Association Test. *Journal of Personality and Social Psychology*, 81, 760-773.

Brown J. (1958). Some tests of the decay theory of immediate memory. Journal of Experimental Psychology,10, 12–21.

Campbell, K. (2010). Event-related potentials as a measure of sleep disturbance: a tutorial review. *Noise and Health,* 12, 137-153.

Chee, M.W.L, Sriram, N., Soon, C.S., & Lee, K.M. (2000). Dorsolateral prefrontal cortex and the implicit association of concepts and attributes. *Neuroreport*, 11, 135-140.

Coates, MA., Mask, L., & Campbell, K. (2005, February). *Window into the biased brain; Frontal lobe activity during the Implicit Association Test.* Poster presented at the annual meeting of the Society For Personality and Social Psychology, New Orleans, LA.

Coates, M.A., & Campbell, K. (2010). Event-related potential measures of processing during an Implicit Association Test . *Neuroreport,* DOI: 10.1097/WNR.0b013e32833f5e7d.

Cunningham, W.A., Johnson,M.K., Raye, C.L., Gatenby, J.C., Gore, J.C., & Banaji, M.R. (2004). Separable neural components in the processing of Black and White faces. *Psychological Science,* 15, 806-813.

Dasgupta, N., McGhee, D.E., Greenwald, A.G., & Banaki, M.R. (2000). Automatic preferences for White Americans: Eliminating the familiarity explanation. *Journal* *of Experimental Social Psychology*, 36, 316-328.

De Houwer, J. (2001). A structural and process analysis of the Implicit Association Test. *Journal of Experimental Social Psychology*, 37, 443-451.

De Houwer, J., Geldof, T., & De Bruycker, E. (2005) The Implicit Association Test as a general measure of similarity. *Canadian Journal of Experimental Psychology*, 59, 228-239.

De Houwer, J., Teige-Mocigemba, S., Spruyt, A. & Moors, A. (2009). Implicit measures; a normative analysis and review. *Psychological Bulletin*, 135, 347-368.

Donchin, E., & Coles, M. G. H. (1988). Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences*, 11, 357–374.

Duncan-Johnson, C.C., & Donchin, E., (1982). The P300 component of the event-related brain potential as an index of information processing. *Biological Psychology*, 14, 1-52.

Duncan-Johnson, C.C., & Kopell, B.S. (1981). The Stroop effect: brain potentials localize the source of interference. *Science*, 214, 938-940.

Egloff B., & Schmukle S.C. (2002). Predictive validity of an Implicit Association Test for assessing anxiety. *Journal of Personality and Social Psychology*, 83, 1441-1455.

Eimer, M. (1993). Effects of attention and stimulus probability on ERPs in a Go/NoGo task. *Psychophysiology,* 35, 128-135.

Escera, C., Alho, K., Schroger, E., & Winkler, I. (2000). Involuntary attention and distractibility as evaluated with event-related brain potentials. *Audiology and Neuro-Otology,* 5, 151, 166.

Escera, C., Yago, E., & Alho, K. (2001). Electrical responses reveal the temporal dynamics of brain events during involuntary switching. *European Journal of Neuroscience*, 14, 877-883.

Fazio, R.H. & Olson, M.A. (2003). Attitudes: Foundations, functions, and consequences. In Hogg, M.A. & Cooper, J. (Eds), *The Sage Handbook of Social Psychology.* Thousand Oaks: Sage Publications.

Francis, W.N., & Kucera, H. (1982). *Frequency Analysis of English Usage: Lexicon and Grammar*. Boston: Houghton Mifflin Company.

Freiberg, K., Tually, K., & Crassini, B. (2001) Use of an auditory looming task to test infants'

sensitivity to sound pressure level as an auditory distance cue. *British Journal of Developmental Psychology, 19*, 1–10.

Folstein, J.R., & Van Petten, C. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: A review. *Psychophysiology*, 45, 152-170.

Fuller, R.L., Luck, S.J., Braun, E.L., Robinson, B.M., McMahon, R.P., & Gold, J.M., Impaired Visual Working Memory Consolidation in Schizophrenia, *Neuropsychology*, 23, 71-80.

Girodo, M., Deck, T.P., & Campbell, K.B. (2002). Event-related potentials reveal the effects of altering personal identity. *Neuroreport*, 13, 1595-1598.

Goffaux, P., Phillips, N.A., Sinai, M., & Pushkar, D. (2006). Behavioural and electrophysiological measures of task switching during single and mixed-task conditions . *Biological Psychology*, 72, 278-290.

Goffaux, P., Phillips, N.A., Sinai, M., & Pushkar, D. (2008). Neurophysiological measures of task-set switching: effects of working memory and aging. *Journal of Gerontology,* 63, 57-66.

Gozzi, M., Raymont, V., Soloman, J., Koenigs, M., & Grafman, J. (2009). Dissociable effects of prefrontal and anterior temporal cortical lesions on stereotypical gender attitudes. *Neuropsychologia,* 47, 2125-2132.

Grafman, J. (1994). Alternative frameworks for the conceptualization of prefrontal lobe functions. In: Boller F., & Grafman, J. (Eds), *Handbook of neuropsychology*. Amsterdam: Elsevier.

Greenwald, A.G., McGhee, D.E., & Schwartz, J.L.K. (1998). Measuring individual difference in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, 74, 1464-1480.

Greenwald, A.G., Nosek, B.A., & Banaji, M.R. (2003). Understanding and using the Implicit Association Test: 1. An improved scoring algorithm. *Journal of Personality and Social Psychology,* 85, 197-216.

Greenwald, A.G., Nosek, B.A., Banaji, M.R., & Klauer, K.C. (2005). Validity of the salience asymmetry interpretation of the Implicit Association Test: comment on Rothermund and Wentura (2004). *Journal of Experimental Psychology: General,* 134, 420-425.

Hammerer, D., Li, S.-C., Muller, V., & Lindenberger, U. (2010). An electrophysiological study of response conflict processing across the lifespan: Assessing the roles of conflict monitoring, cue utilization, response anticipation, and response suppression . *Neuropsychologia*, 48, 3305-3316.

Hauk, O., Davis, M.H., Ford, M., Pulvermüller, F., and Marslen-Wilson, W.D. (2006). The time course of visual word recognition as revealed by linear regression analysis of ERP data. *NeuroImage*, 30, 1383-1400.

Hauk, O., Patterson, K., Woollams, A., Watling, L., Pulvermüller F., Rogers, T.T. (2006) [Q:] When Would You Prefer a SOSSAGE to a SAUSAGE? [A:] At about 100 msec. ERP Correlates of Orthographic Typicality and Lexicality in Written Word Recognition. *Journal of Cognitive Neuroscience*, 18, 818-832.

He, Y., Johnson, M.K., Dovidio, J.F., & McCarthy, G. (2009). The relation between race-related implicit associations and scalp-recorded neural activity evoked by faces from different races. *Social Neuroscience*, 4, 426-442.

Hoffman, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2005). A meta-analysis on the correlation between the Implicit Association Test and explicit self-report measures. *Personality and Social Psychology Bulletin*, 31, 1369-1385.

Hugenberg, K., & Bodenhausen, G.V. (2003). Facing prejudice: Implicit prejudice and perception of facial threat. *Psychological Science,* 14, 640-643.

Hurtado, E., Haye, A., Gonzalez, R., Manes, F. & Ibanez, A. (2009). Contextual blending of ingroup/outgroup face stimuli and word valence: LPP modulation and convergence of measures. *BMC Neuroscience,* 10.

Ito, T.A., & Bartholow, B.D. (2009). The neural correlates of race. *Trends in Cognitive Science,* 13, 524-531.

Johnson, R. (1986). A triarchic model of P300 amplitude. *Psychophysiology*, 23, 367-384.

Johnson, R., & Donchin, E. (1985). Second thoughts; Multiple P300s elicited by a single stimulus. *Psychophysiology*, 22, 182-194.

Johnson, R., & Donchin, E. (1978). On how P300 amplitude varies with the utility of the eliciting stimulus.

Kieffaber, P.D., & Hetrick, W.P. (2005). Event-related potential correlates of task switching and switch costs . *Psychophysiology*, 42, 56-71.

Kim, S., Kim, M., & Chun, M.M. (2005). Concurrent working memory load can reduce distraction. *Proceedings of the National Academy of Science*, 102, 16524-16529.

Klauer, C.K. & Mierke, J. (2005). Task-set inertia, attitude accessibility, and compatibility-order effects: New evidence for a task set switching account of the Implicit Association Test effect. *Personality and Social Psychology Bulletin*, 31, 208-217.

Klauer, C.K., Schmitz, F., Teige-Mocigemba, S., & Voss, A. (2009). Understanding the role of executive control processes in the Implicit Association Test: Why flexible people have small IAT effects. *The Quarterly Journal of Experimental Psychology*, DOI: 10.1080/17470210903076826.

Kok, A. (1997). Event-related-potential (ERP) reflections of mental resources: a review and synthesis. *Biological Psychology*, 45, 19-56.

Kok, A. (2001). On the utility of P3 amplitude as a measure of processing capacity. *Psychophysiology*, 38, 557-577.

Koole, S., Govorun, O., Cheng, C.M., & Gallucci, M. (2009). Pulling yourself together: Meditation promotes congruence between implicit and explicit self-esteem. *Journal of Experimental Social Psychology*, 45, 1220-1226.

Kutas, M., McCarthy, G., & Donchin, E. (1977) Augmenting mental chronometry; the P300 as a measure of stimulus evaluation time. *Science,* 197, 792-795.

LaPierre, R.T. (1934) Attitudes vs. Actions. *Social Cognition*, 13, 230-237.

Lavie, N. (2010). Attention, distraction, and cognitive control under load. *Current Directions in Psychological Science*, 19, 143-148.

Lavie, N. & de Fockert, J. (2005). The role of working memory in attention capture.

Psychonomic Bulletin and Review, 12, 669-674.

Liotti, M., Woldorff, M.G., Perez, R., & Mayberg, H.S. (2000). An ERP study of the temporal course of the Stroop color-word interference effect. *Neuropsychologia*, 38, 701-711.

Macdonald, M., Jamshidi, P., & Campbell, K. (2008). Infrequent increases in stimulus intensity may

interrupt central executive functioning during Rapid Eye Movement sleep. *Neuroreport*, *19*, 309-13.

Maier, J. & Ghazanfar A. (2007). Looming biases in monkey auditory cortex. *The Journal of Neuroscience, 27*, 4093–4100.

Maison, D, Greenwald, AG. & Bruin, R.H. (2004). Predictive Validity of the Implicit Association Test in Studies of Brands, Consumer Attitudes, and Behavior. *Journal of Consumer Psychology*, 14, 405-415.

McCarthy, C., & Donchin, E. (1981). A metric for thought: a comparison of P300 latency and reaction time. *Science*, 211, 77-80.

McConnell, A.R., & Leibold, J.M. (2001). Relations among the Implicit Association Test, discriminator behaviour, and explicit measures of racial attitudes. *Journal of Experimental Social Psychology*, 37, 435-442.

McDermott, L.M., & Ebmeier, K.P. (2009). A meta-analysis of depression severity and cognitive function. *Journal of Affective Disorders*, 119, 1-8.

Mierke, J., & Klauer, K.C. (2001). Implicit association measurement with the IAT: Evidence for effects of executive control processes. *Zeitschrift fur Experimentelle Psychologie*, 48, 107-122.

Mierke, J. & Klauer, K.C. (2003). Method specific variance in the Implicit Association Test.

Journal of Personality and Social Psychology, 85, 1180-1192.

Milne, E., & Grafman, J. (2001). Ventromedial prefrontal cortex lesions in humans eliminate implicit gender stereotyping. *Journal of Neuroscience*, 21:RC150, 1-6.

Mitchell, J.P., Ames, D.L., Jenkins, A.C., & Banaji, M.R. (2009). Neural correlates of stereotype activation. *Journal of Cognitive Neuroscience, 21, 594-604.*

Muller-Gass, A., Gonthier, I., Desrochers, A., & Campbell, K.B. (2000). Multiple P3 evidence of a two-stage process in word gender decision. *Neuroreport*, 11, 3527-3531.

Muller-Gass, A., Stelmack, R., & Campbell, K. (2006). The effect of visual task difficulty and attentional direction on the detection of acoustic change as indexed by the Mismatch Negativity. *Brain Research,* 1078, 112-130.

Muller-Gass, A., Macdonald, M., Schroger, E., Sculthorpe, L., & Campbell, K.B. (2007). Evidence for the auditory P3a reflecting an automatic process: elicitation during highly-focused continuous visual attention. *Brain Research, 1170*, 71-78.

Munka, L., & Berti, S. (2006). Examining task-dependencies of different attentional processes as reflected in the P3a and reorienting negativity components of the human event-related potential. *Neuroscience Letters*, 396, 177-181.

Murray, L.J., & Ranganath, C. (2007). The Dorsolateral Prefrontal Cortex Contributes to Successful Relational Memory Encoding. *The Journal of Neuroscience,* 27, 5515-5522.

Näätänen R, Kujala T, Winkler I. (2010). Auditory processing that leads to conscious perception: A unique window to central auditory processing opened by the mismatch negativity and related responses. *Psychophysiology*, DOI: 10.1111/j.1469-

8986.2010.01114.x.

Nosek, B.A., & Banaji, M.R. (2001). The Go/No-Go Association Task. *Social Cognition*, 19, 625-664.

Nosek, B.A., Greenwald, A.G., & Banaji, M.R. (2005). Understanding and using the Implicit Association Test: II. Method variables and construct validity. *Personality and Social Psychology Bulletin*, 31, 166-180.

Nosek, B.A., Greenwald, A.G., & Banaji, M.B. (2007). The Implicit Association Test at age 7: methodological and conceptual review (p. 265-292). In J.A. Bargh (Ed.), *Automatic processes in social thinking and behavior.* Psychology Press.

Pashler, H., Johnston, J.C., & Ruthruff, E. (2001). Attention and performance. *Annual Review of Psychology,* 52, 629-651.

Petrides, M. (2000). The role of the mid-dorsolateral prefrontal cortex in working memory. *Experimental Brain Research*, 133, 44-54.

Pfefferbaum, A., Ford, J.M., Weller, B.J., & Kopell, B.S. (1985). ERPs to response production and inhibition. *Electroencephalography and Clinical Neurophysiology,* 60, 423-434.

Phelps, E.A., O'Connor, K.J., Cunningham, W.A., Funayama, E.S., Gatenby, J.C., Gore, J.C., Banaji, M.R. (2000). Performance on indirect measures of race evaluation predicts amygdala activation. *Journal of Cognitive Neuroscience*, 12, 729-738.

Plante, E.A., Devine, P.G., Cox, W.T.L., Columb, C., Miller, S.L., Goplen, J., & Peruche, B.M. (2009). The Obama effect; decreasing implicit prejudice and stereotyping. *Journal of Experimental Social Psychology*, 45, 961-964.

Poldrack , RA., Wagner, AD., & Mitchell, J.P. (2008) Contributions of functional neuroimaging to the study of social cognition. *Current Directions in Psychological Science*, 17, 142-146.

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, 118, 2128-2148.

Postle, B.R. (2006). Working memory as an emergent property of the mind and brain. *Neuroscience*, 139, 23-38.

Richeson, J.A., & Shelton, J.N. (2003). When prejudice does not pay; Effects of interracial contact on executive functions. *Psychological Science*, 14, 287-290.

Rinne, T., Särkkäc, A., Degerman, A, Schröger, E., & Alho, K. (2006). Two separate mechanisms underlie auditory change detection and involuntary control of attention. *Brain Research,* 1077, 135-143.

Rothermund, K., & Wentura, D. (2001). Figure-Ground asymmetries in the Implicit Association Test (IAT). *Zeitschrift fur Experimentelle Psychologie*, 48, 94-106.

Rothermund, K., & Wentura, D. (2004). Underlying processes in the Implicit Association Test: Dissociating salience from associations. *Journal of Experimental Psychology: General*, 133, 139-165.

Rudman, L.A., & Glick, P. (2001). Prescriptive gender stereotypes and backlash towards agentic women. *Journal of Social Issues*, 57, 743-762.

Rushworth, M.F.S., Passingham, R.E., & Nobre, A.C. (2002). Components of switching intentional set. *Journal of Cognitive Neuroscience*, 14, 1139-1150.

SanMiguel, I., Corral, M.-J., & Escera, C. (2008). When loading working memory reduces distraction: behavioral and electrophysiological evidence from an auditory-visual distraction paradigm. *Journal of Cognitive Neuroscience*, 20, 1131-1145.

Schuman, H., Steeh, C., & Bobo, L. (1997). *Racial attitudes in America: Trends and interpretations*. Cambridge: Harvard University Press.

Sereno, S.C., Rayner, K., & Posner, M.I. (1998). Establishing a time-line of word recognition: evidence from eye movements and event-related potentials. *Neuroreport*, 9, 2195-2200.

Squires, K.C., Wickens, C., Squires, N., & Donchin, F. (1976) The effect of stimulus sequence on the waveform of the cortical event-related potential. *Science*, 193, 1142-1146.

Steffens, M.C., & Buchner, A. (2003). Implicit Association Test; Separating transsituationally stable and variable components of attitudes towards gay men. *Experimental Psychology*, 50, 33-48.

Stroop, J.R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology,* 18, 643-662.

Stuss, D.T., Gow, C.A., & Hetherington, C.R. (1992). “No longer Gage: frontal lobe dysfunction and emotional changes. *Journal of Consulting and Clinical Psychology*, 60, 349-359.

Sussman, E., Winkler, I., & Schroger, E. (2003). Top-down control over involuntary attention switching in the auditory modality . *Psychonomic Bulletin and Review*, 10, 630-637.

Swainson, R., Cunnington, R., Jackson, G.M., Rorden, C., Peters, A.M., Morris, P.G. & Jacks , S.R. (2003). Cognitive control mechanisms revealed by ERP and fMRI: evidence from repeated task-switching . *Journal of Cognitive Neuroscience*, 15, 785-799.

Thush, C, Wiers, R.W., Ames, S.L., Grenard, J.L., Sussman, S., & Stacy , A.W. (2008). Interactions between implicit and explicit cognition and working memory capacity in the prediction of alcohol use in at-risk adolescents . *Drug and Alcohol Dependence*, 94, 116-124.

Van Overwalle, F. (2009). Social cognition and the brain: A meta-analysis. *Human Brain Mapping,*30, 829-858.

Vergeler, R. (1997). On the utility of P3 latency as an index of mental chronometry. *Psychophysiology,* 34, 131-156.

Verleger, R., Japkowski, P., & Wascher, E. (2005). Evidence for an integrated role of P3b in linking reaction to perception. *Journal of psychophysiology,* 19, 165-181.

Vul, E., Harris, C., Winkielman, P., & Pashler, H. (2009). Puzzlingly high correlations in fMRI studies of emotion, personality, and social cognition. *Perspectives in Psychological Science,* 4, 274-290.

Wechsler (2009). Wechsler Memory Scale – Fourth Edition. Pearson: San Antonio.

Williams J.K, & Themansen, J.R. (2010). Neural correlates of the Implicit Association Test: evidence for semantic and emotional processing. *Social Cognitive and Affective Neuroscience,* 2010, 1-9.

Woestenberg J, Verbatan M, & Slangen J. (1983). The removal of the eye-movement artifact from the EEG by regression analysis in the frequency domain. *Biological Psychology*; 16**,** 127-147.

Table 1

Schematic Description of the IAT

Block Name	Stimulus	Response Required	Total Number of Trials
Target Acquisition	Weapon	L	25
	Music	R	25
Attribute Acquisition	Bad	L	25
	Good	R	25
Congruent	Weapon	L	50
	Bad	L	50
	Music	R	50
	Good	R	50
Target Reverse	Music	L	25
	Weapon	R	25
Incongruent	Music	L	50
	Good	R	50
	Weapon	R	50
	Bad	L	50

Note. "L" indicates that the participant is required to press the left-hand key, and "R" indicates the right-hand key.

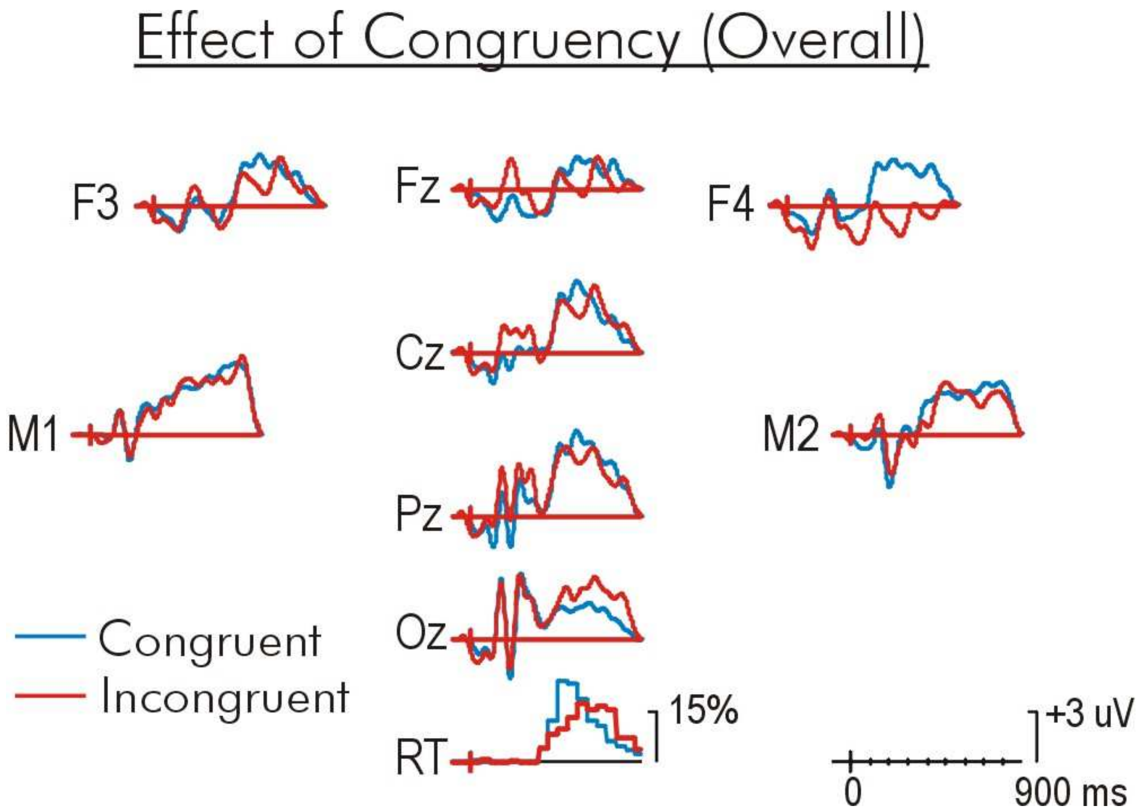

Figure 1. Grand average of Congruent and Incongruent conditions.

All trials within the Congruent and Incongruent blocks. At the bottom of the Figure, 50ms RT histograms are provided. RT is faster in the Congruent trials than in Incongruent trials. The ERPs at the different electrode sites are also traced. Positivity of the scalp is indicated by an upward deflection. A large positive deflection is apparent at 400-800ms after stimulus presentation at central (Cz) and parietal (Pz) sites. The initial deflection occurs at about the same time in the Congruent and Incongruent conditions. However, a second positive peak occurs somewhat earlier for the Congruent condition than for Incongruent condition.

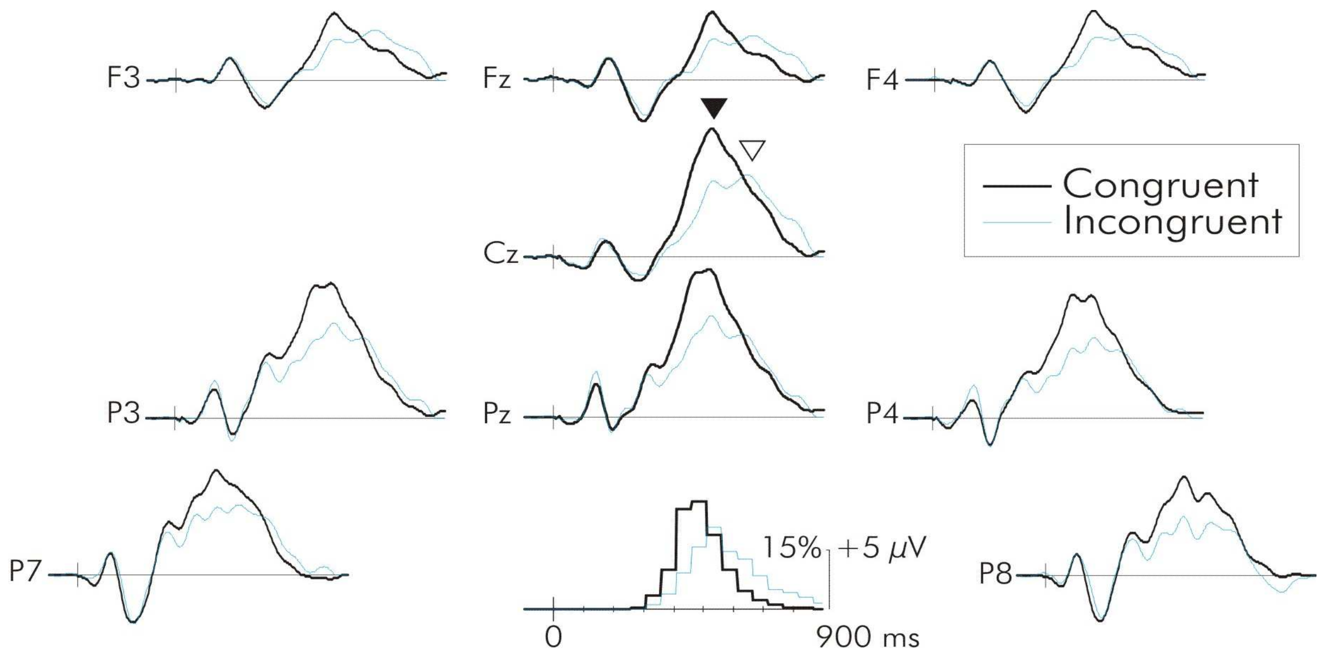

Figure 2. Grand average – Experiment 1

Grand-average (average of all participants' averages) waveforms during the Congruent and Incongruent conditions. At the bottom of the figure, the RT histogram bins are displayed. Bin width is 50 ms. The amplitude of the bin corresponds to the percentage of trials falling within it. Note that the overall RT is delayed in the Incongruent trials. A late positive ERP component, the P300 (indicated by the closed arrow) was elicited by the words to-be-categorized in both the Congruent and Incongruent. The latency to this peak did not however differ between the Congruent and Incongruent conditions. In the Incongruent condition, the words elicited a second positivity (indicated by the open arrow). The ERP data suggest that classification of the words requires a single decision when presented in the Congruent condition but a double decision when presented in the Incongruent condition.

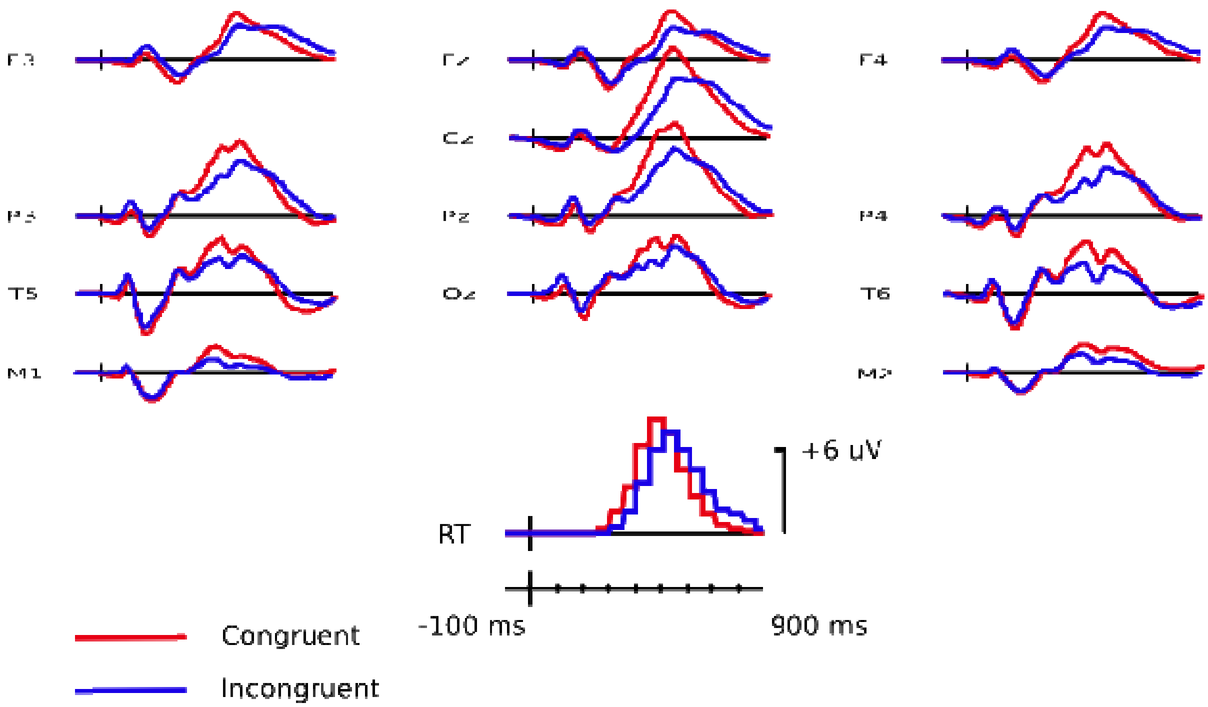

Figure 3. Experiment 2 – Grand averages for all subjects in Congruent and Incongruent conditions.

A distinct P300 was observed in both the Congruent and the Incongruent conditions. With more subjects than in the previous study, the second positive peak in the Incongruent condition is not evident, although this component does appear to be slightly longer lasting than in the Congruent condition.

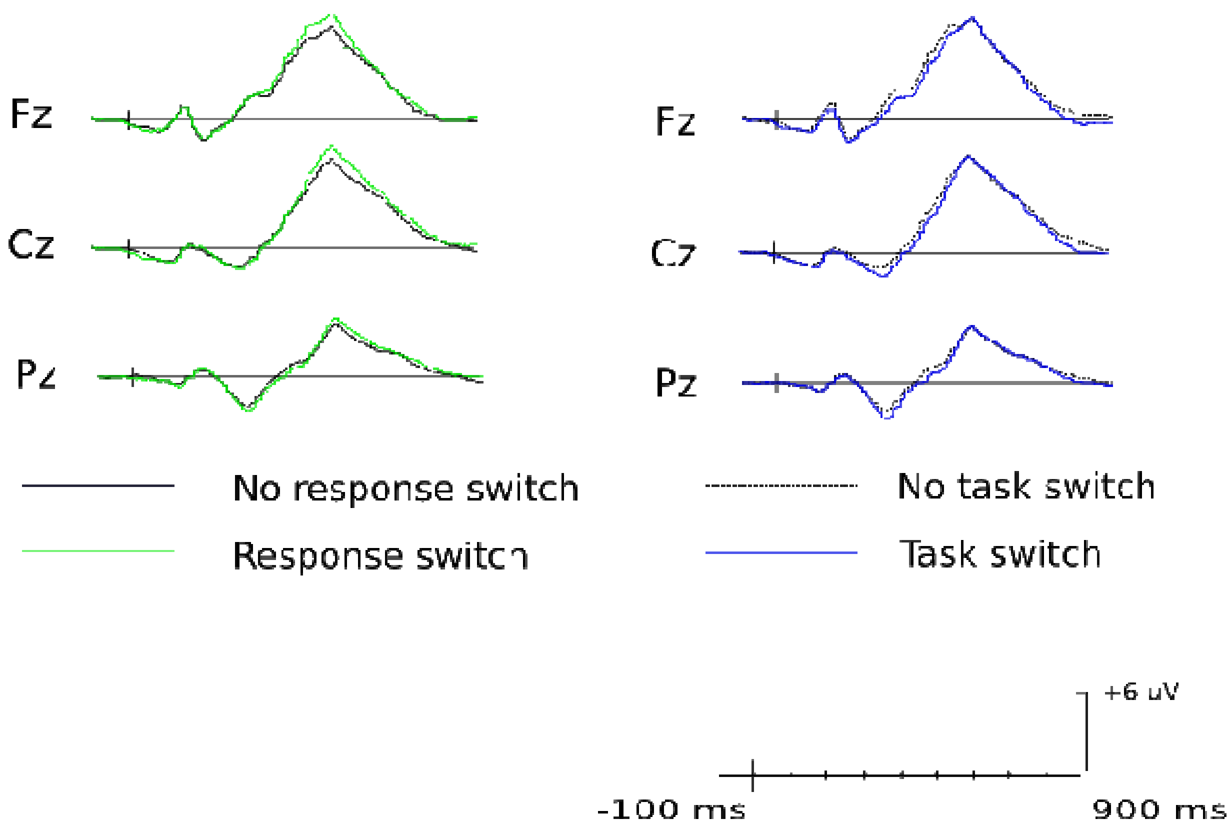

Figure 4. Experiment 2 – Effect of task set switching in the Congruent condition. Within the Congruent condition, all trials elicited a very similar amplitude and latency P300 component, regardless of whether a trial technically required a task set switch or response switch. This finding is consistent with the task set switching account of the IAT, which argues that all trials of the Congruent conditions are processed similarly.

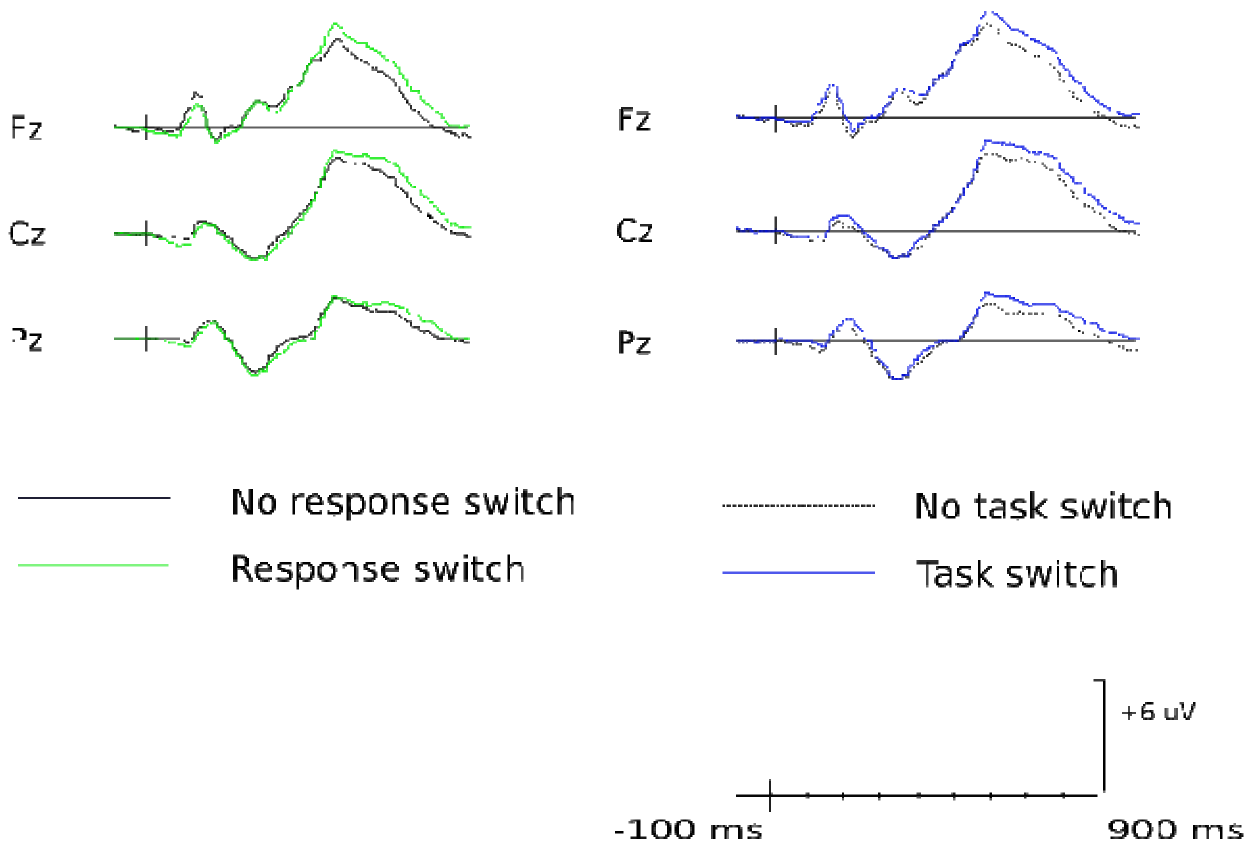

Figure 5. Experiment 2 – Effect of task set switching in the Incongruent condition. Within the Incongruent condition, trials containing a switch (response switch, task set switch, or both) elicited significantly greater amplitude P300 than trials that did not require switching. This finding is inconsistent with the task set switching account, which would predict less equivocation on no-switch trials (ie: these trials should be processed similarly to trials of the Congruent condition). The effect of local probability (sequential expectancy) may explain the larger P300 amplitude in trials where a switch is required.

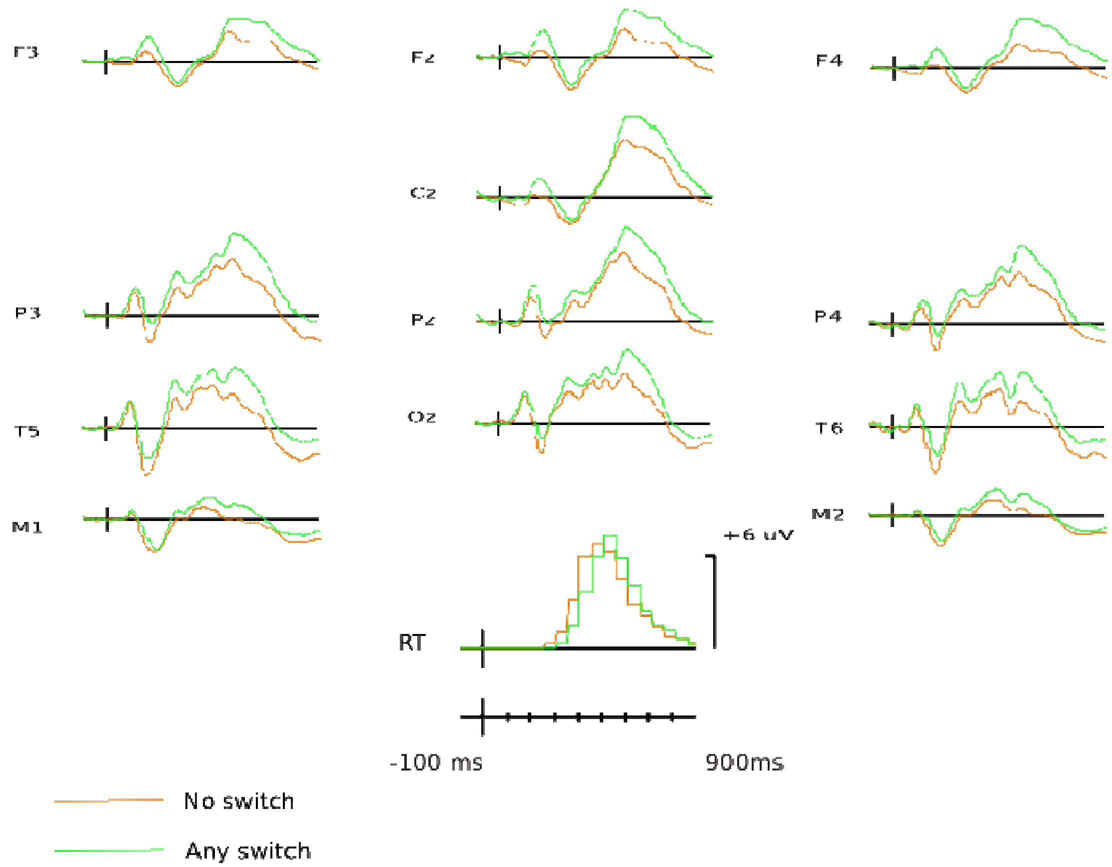

Figure 6. Experiment 2 – Incongruent condition, switch versus no-switch trials.

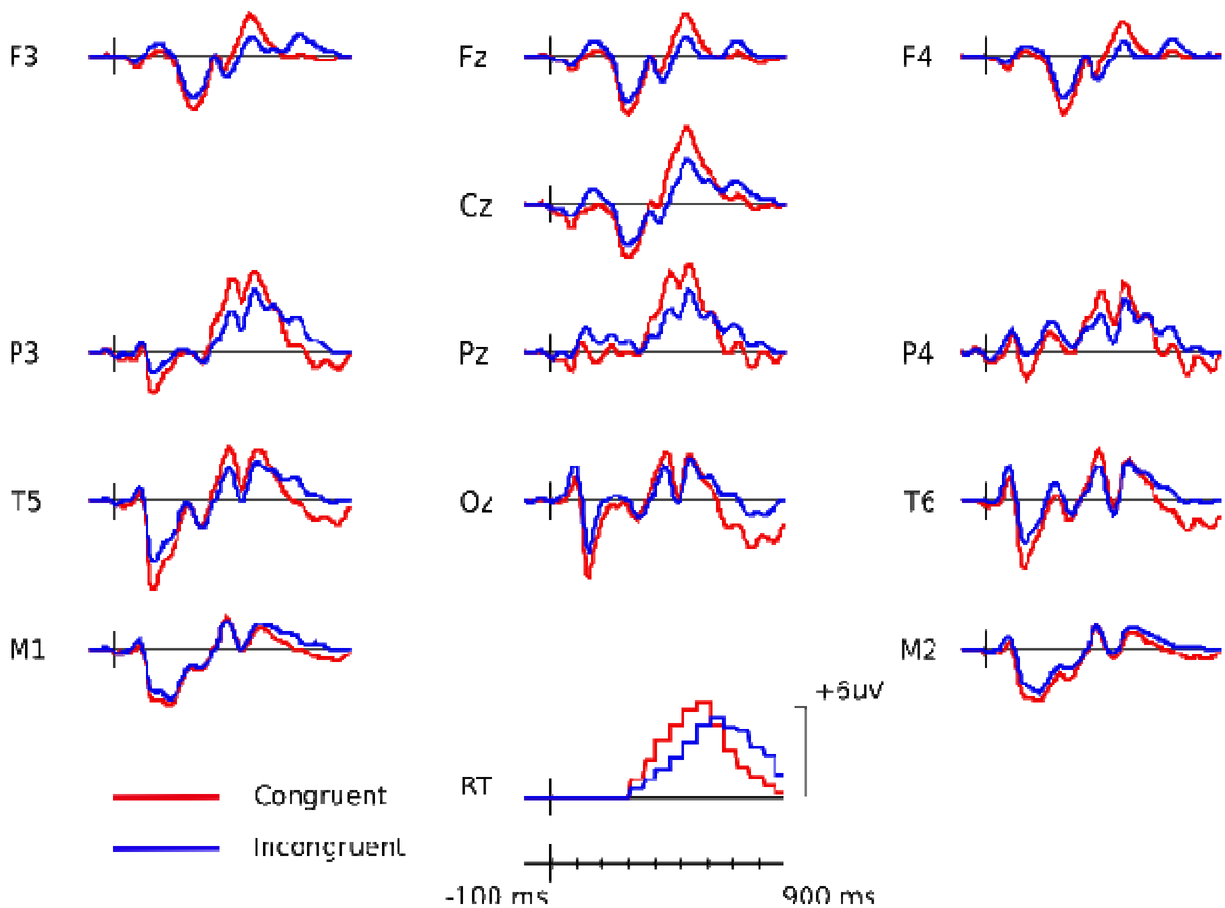

Figure 7. Experiment 3 – Grand average of Congruent and Incongruent conditions for all trial types.

The overall grand averages for all trials in the Congruent and Incongruent conditions were similar to those observed in previous experiments. A double peaked P300 was visible at Cz and Pz in both the Congruent and Incongruent conditions, occurring at approximately 350 ms and 450 ms in both conditions. In the Incongruent condition, a slower return to baseline following the P300 peak was apparent, as was observed in previous experiments.

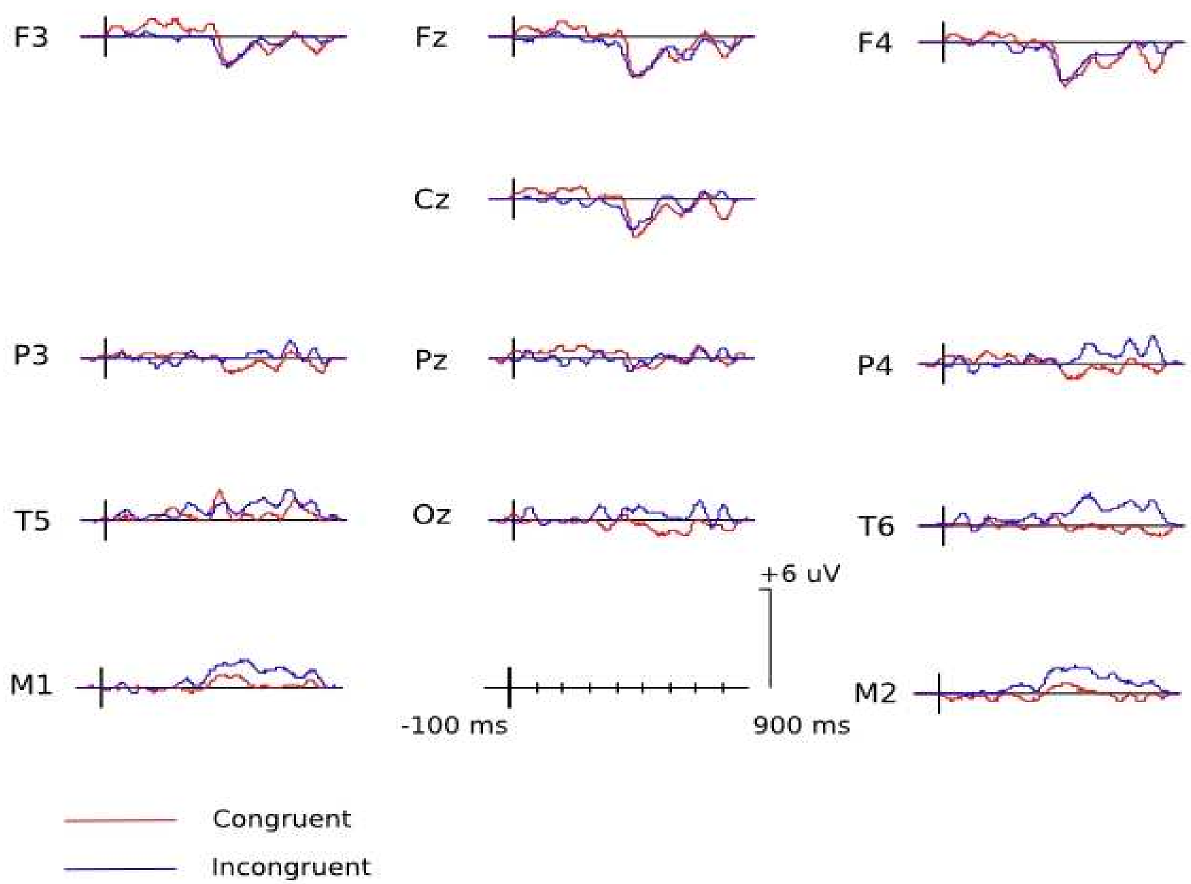

Figure 8. Experiment 3 – Difference waveforms (deviant minus standard auditory stimuli) for Congruent and Incongruent conditions.

When trials containing standard auditory stimuli are subtracted from trials containing deviant auditory stimuli, the N1/MMN component is visible at Fz and Cz. This negativity begins approximately 100 ms after onset of the auditory stimulus, which was approximately 400 ms after onset of visual stimulus. There is no apparent difference in the amplitude of this negativity at the critical frontal and central electrode sites between the Congruent and Incongruent conditions.

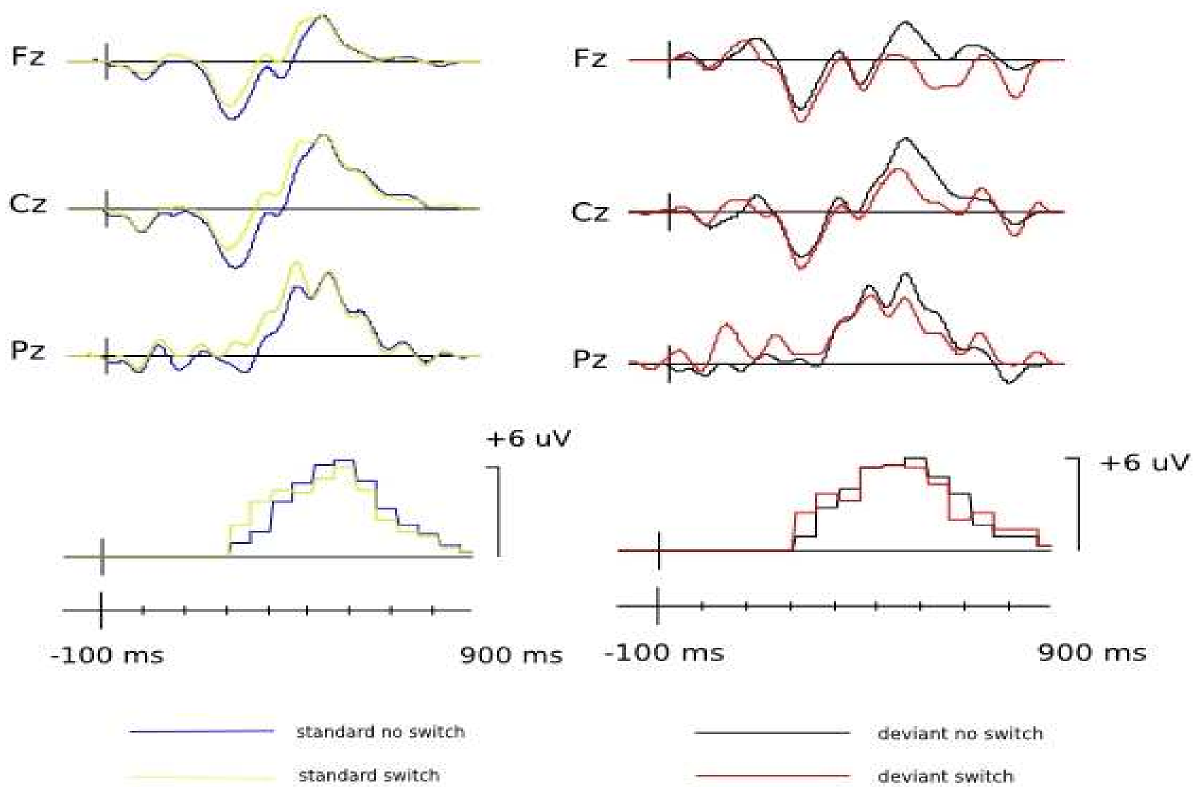

Figure 9. Experiment 3 – Grand averages of Congruent condition, divided according to the type of auditory stimulus (standards and deviants).

On trials of the Congruent condition that are preceded by a standard tone (left side of figure), there appears to be slightly greater negativity of N2 on no-switch trials than switch trials. Very little difference in P300 amplitude is apparent between switch and no-switch trials. On deviant trials of the Congruent condition, no differences in N2 amplitude are visible when comparing switch and no-switch trials. No-switch trials elicited a larger amplitude P300 than switch trials at frontal and central sites.

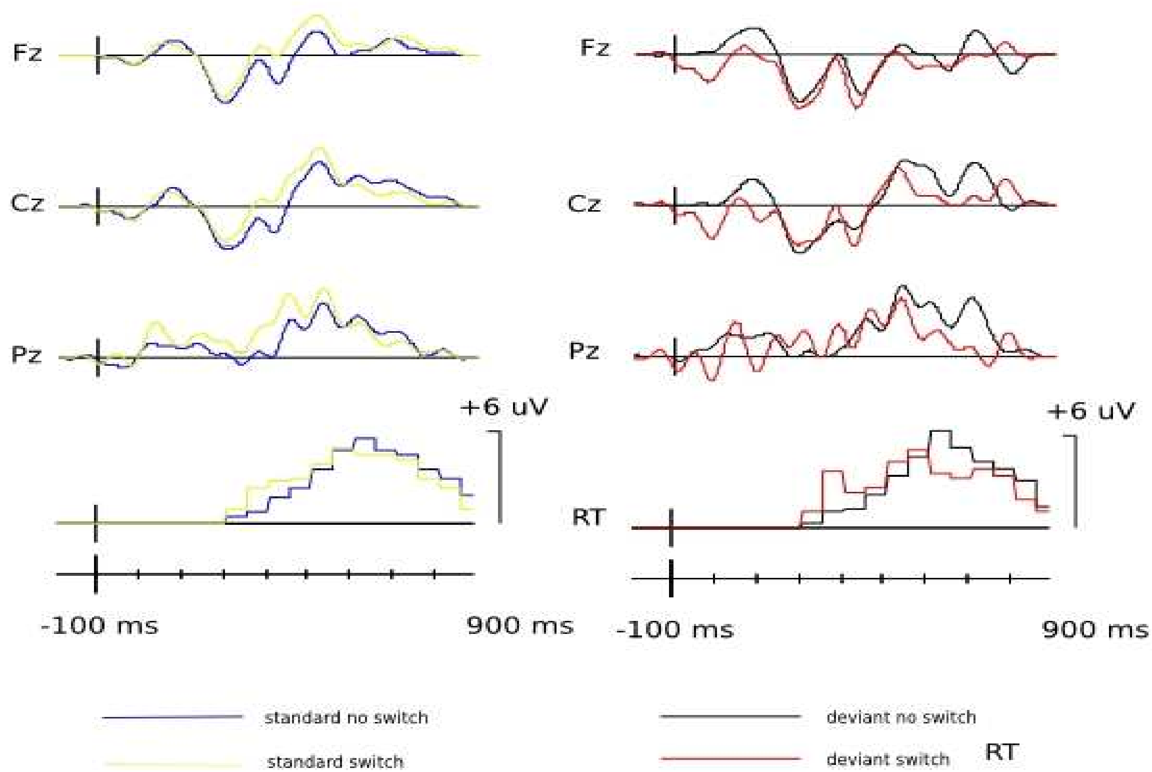

Figure 10. Experiment 3 – Grand averages of Incongruent condition, divided according to the type of auditory stimulus (standards and deviants).

Similar to the Congruent condition, on trials of the Incongruent condition on which standard tones were presented there was a slightly greater negativity of N2 for no-switch trials. P300 amplitude was also somewhat smaller on no-switch trials. A double peaked positivity was visible at Pz for both switch and no-switch trials. In the Incongruent condition when trials are preceded by a deviant tone, differences in N2 amplitude were not apparent, possibly due to a relatively high level of background noise. There was also little visible difference in P300 amplitude between switch and no-switch trials. However, the P300 peak elicited on no-switch trials appears to be somewhat longer lasting at Cz than on switch trials.

Appendix A

Stimuli used in Pilot Study

Good words: Caress, Freedom, Health, Love, Peace, Cheer, Friend, Heaven, Loyal, Pleasure, Diamond, Gentle, Honest, Lucky, Rainbow, Diploma, Gift, Honor, Miracle, Sunrise, Family, Happy, Laughter, Paradise, Vacation

Bad words: Abuse, Crash, Filth, Murder, Sickness, Accident, Death, Grief, Poison, Stink, Assault, Disaster, Hatred, Pollute, Tragedy, Pain, Divorce, Jail, Poverty, Ugly, Cancer, Evil, Kill, Rotten, Agony

Weapons: Arrow, Club, Gun, Missile, Spear, Axe, Dagger, Harpoon, Pistol, Sword, Blade, Dynamite, Hatchet, Rifle, Tank, Lance, Firearm, Knife, Shotgun, Tear Gas, Cannon, Grenade, Slingshot, Whip, Mace

Musical Instruments: Bagpipe, Cello, Guitar, Lute, Trombone, Banjo, Clarinet, Harmonica, Mandolin, Trumpet, Bassoon, Drum, Harp, Oboe, Tuba, Bell, Fiddle, Harpsichord, Piano, Violin, Bongo, Flute, Horn, Saxophone, Viola

Words were selected from those used by Greenwald and colleagues (1998) in the original article on the IAT.

Appendix B

Stimuli used in Experiments 1, 2, and 3

Good: Freedom (133), Family (405), Peace (198), Love (145), Health (105)

Bad: Abuse (30), Crash (40), Sickness (57), Accident (41), Hatred (20)

Weapons: Gun (142), Knife (86), Rifle (87), Tank (30), Pistol (31)

Musical Instruments: Violin (13), Horn (33), Guitar (22), Drum (26), Piano (39)

Note: The number in parenthesis is the frequency of usage of that word or very similar forms of that word (for example, “sickness” includes the word frequency of the word “sick”) in a representative sample of English text (Francis & Kucera, 1982). The authors who compiled this word frequency list indicate that low frequency words are considered to be any word with a frequency of less than 5. Therefore, all the words used in the current experiments are considered to be common words.

Acknowledgements

I would like to thank my committee members Dr. Vernor Knott, Dr. Francine Tougas, and Dr. Charles Collin for their time and feedback. Special thanks to my supervisor Dr. Ken Campbell for his support, time, and expertise. I would also like to acknowledge the assistance of the many administrative staff at the University of Ottawa who helped me throughout graduate school, particularly Mireille Coté, Martine Raymond, and Dr. John Hunsley. To WRB, thank you for your support during the early years of my doctoral program. Invaluable encouragement and advice on statistics were provided by my partner, Dr. Gillian Stanley. And finally, I would like to thank my parents, Ken and Cathy Coates, and grandparents, Earl and Anne Shepherd and Marge Coates, for the great help that you have provided.